

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2004, 22:13:32 ; Search time 156.647 Seconds  
(without alignments)  
7896.639 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 caggagagccaccagccctca.....aaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY\_NUC

Gap 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/prodata/2/ina/ECTUS\_COMB.seq:\*
- 6: /cgn2\_6/prodata/2/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	48.2	2.2	8607	4 US-10-204-708-72	Sequence 72, Appl
2	47.4	2.1	1937	4 US-09-647-143-1	Sequence 1, Appl
3	45.8	2.1	1048	4 US-09-489-847-38	Sequence 38, Appl
C 4	45.4	2.0	615	4 US-09-252-991A-10075	Sequence 10075, A
5	45.4	2.0	1126	4 US-09-252-991A-9937	Sequence 9937, Ap
6	45.4	2.0	1926	4 US-09-252-991A-9758	Sequence 9758, Ap
C 7	43.2	1.9	2823	4 US-09-252-991A-7100	Sequence 7100, Ap
8	43.2	1.9	3456	4 US-09-252-991A-7025	Sequence 7025, Ap
9	42.6	1.9	558	4 US-09-252-991A-6738	Sequence 6738, Ap
10	42.6	1.9	787	4 US-09-621-976-1878	Sequence 1878, Ap
11	42.6	1.9	2346	4 US-09-252-991A-6782	Sequence 6782, Ap
12	42.4	1.9	630	4 US-09-252-991A-3386	Sequence 3386, Ap
13	42.4	1.9	1647	4 US-09-252-991A-3344	Sequence 3344, Ap
C 14	42.2	1.9	1089	4 US-09-796-766-13	Sequence 13, Appl
15	42.2	1.9	804	4 US-09-252-991A-14038	Sequence 14038, A
C 16	42.2	1.9	807	4 US-09-252-991A-14227	Sequence 14227, A
17	42.2	1.9	2083	4 US-09-716-129-41	Sequence 41, Appl
18	41.8	1.9	159	4 US-09-621-976-17448	Sequence 17448, A
C 19	41.8	1.9	1653	4 US-09-252-991A-10546	Sequence 10546, A
20	41.8	1.9	1701	4 US-09-252-991A-10316	Sequence 10316, A
21	41.6	1.9	1212	4 US-09-182-145-34	Sequence 34, Appl
C 22	41.6	1.9	1212	4 US-09-182-145-35	Sequence 35, Appl
23	41.4	1.9	276	4 US-09-621-976-16611	Sequence 16611, A
24	41.4	1.9	2394	4 US-09-800-729-33	Sequence 33, Appl
25	41.2	1.8	244	4 US-09-621-976-484	Sequence 484, Ap
26	41.2	1.8	298	4 US-09-621-976-3871	Sequence 3871, Ap
27	41.2	1.8	1161	1 US-08-086-439C-2	Sequence 2, Appl

28	41.2	1.8	1161	1 US-08-434-877-2	Sequence 2, Appl
29	41.2	1.8	1367	3 US-08-475-742-3	Sequence 3, Appl
30	41.2	1.8	1367	4 US-08-261-293-3	Sequence 1, Appl
31	41.2	1.8	1370	1 US-08-056-051-1	Sequence 1, Appl
32	41.2	1.8	1370	1 US-07-928-611-17	Sequence 17, Appl
33	41.2	1.8	1370	2 US-08-487-811A-17	Sequence 17, Appl
34	41.2	1.8	1370	3 US-09-060-694-17	Sequence 17, Appl
35	41.2	1.8	1370	4 US-09-378-074-17	Sequence 17, Appl
36	41.2	1.8	1370	5 PCR-US93-07370-17	Sequence 17, Appl
37	41.2	1.8	1466	1 US-08-056-051-3	Sequence 3, Appl
38	41.2	1.8	1466	1 US-07-928-611-19	Sequence 19, Appl
39	41.2	1.8	1466	2 US-08-487-811A-19	Sequence 19, Appl
40	41.2	1.8	1466	3 US-09-060-694-19	Sequence 19, Appl
41	41.2	1.8	1466	4 US-09-378-074-19	Sequence 19, Appl
42	41.2	1.8	1466	5 PCR-US93-07370-19	Sequence 19, Appl
43	41.2	1.8	1504	4 US-09-016-434-1276	Sequence 1276, Ap
44	41.2	1.8	1507	4 US-09-453-323-1	Sequence 1, Appl
45	41.2	1.8	1610	1 US-08-056-051-5	Sequence 5, Appl

## ALIGNMENTS

RESULT 1  
US-10-204-708-72/c  
; Sequence 72, Application US/10204708  
; Patent No. 6677731  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBERCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; TITLE OF INVENTION: By Assessing DNA Methylation  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/204,708  
; PRIOR FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 72  
; LENGTH: 8607  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-204-708-72

Query Match 2.2%; Score 48.2; DB 4; Length 8607;  
Best Local Similarity 54.1%; Pred. No. 0.022;  
Matches 98; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY	2046	TTCAATAAAGCTGTAAGCTGACGACCACTTGAATTAAGTCACTGAT	2105
DB	4244	TTCAATTAATCACTTTTAAAAAATCTTAAATTAATTAATTAATTAAT	4185
QY	2106	TCAGAGTCGGGAACCTTAGTTCATTCGAATTCAGACGACCACTTGAAT	2165
DB	4184	TCATCACTCACTCAAAAAATCTTAATTCACAAATTAATTAATTAATTA	4125
QY	2166	CCAATTAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTA	2225
DB	4134	ACCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	4065
QY	2226	A 2226	

DB 4064 A 4064

## RESULT 2

US-09-647-143-1  
; Sequence 1, Application US/09647143  
; Patent No. 6680196  
; GENERAL INFORMATION:  
; APPLICANT: Batria, Surinder K.  
; APPLICANT: Hollingsworth, Michael A.  
; APPLICANT: University of Nebraska Board of Regents  
; TITLE OF INVENTION: No. 6680196: Gene That is Amplified and  
; TITLE OF INVENTION: Overexpressed in Cancer and Methods of Use Thereof  
; FILE REFERENCE: UNMC63121  
; CURRENT APPLICATION NUMBER: US/09/647,143  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US99/06633  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/079,649  
; PRIOR FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 1937  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-647-143-1

Query Match 2.1%; Score 47.4; DB 4; Length 1937;  
Best Local Similarity 83.1%; Pred. No. 0.014; Indels 0; Gaps 0;  
Matches 54; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2165 CCCAACTAATGATTTAATTAATCAATATCTGTTAAAAA 2224  
DB 1854 CCCAACCTTGCTGTATTAATCAATCTCTTTAAAAA 1913  
QY 2225 AAAAA 2229  
DB 1914 AAAAA 1918

## RESULT 3

US-09-489-847-38  
; Sequence 38, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 1048  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-489-847-38

Query Match 2.1%; Score 45.8; DB 4; Length 1048;  
Best Local Similarity 56.2%; Pred. No. 0.027;

Matches 86; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 2077 ATTGAAACATTAACTGAGCTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGAA 2136  
DB 804 ACTGTCACTACTACCTATCTGTCGACAGGATTCAGAAATCTGGAACCTTA 863  
QY 2137 TCCAGACGACGACGACCTAGTATCTGCCCCAACTAAATGATTAATTAATCAATATC 2196  
DB 864 TTAGCTGTGCGCAATATTACTCTCTTATGATCAATCAATTAATTAATTCCTTA 923  
QY 2197 TCGTTAAAAA 2229  
DB 924 TTAAGAAAAA 956

## RESULT 4

US-09-252-991A-10075/c  
; Sequence 10075, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 10075  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-10075

Query Match 2.0%; Score 45.4; DB 4; Length 615;  
Best Local Similarity 48.6%; Pred. No. 0.025; Indels 0; Gaps 0;  
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 318 TACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTTTACGTGAGCGCTGTGACC 377  
DB 450 TTCATGCTGCTGATGCTGTTCGCGCTGTCCGCCCTGGGTCGCCCTTACGATCCG 391  
QY 378 CGCTGTGAAACGATGAGAACTGCGGCGCCGACCGCTCATGAGCGTGTGTCG 437  
DB 390 AGCGAGCAAGTCCGGAACTTCTGCTGAACCCCGCATCTGCTGGAAGGGGCGAAGCG 331  
QY 438 GGTTCGTGGAAGGCAAGGACGAGCAAGCGCGCTGTGCGGCGACGCTATCCGTAC 497  
DB 330 CGCTTCTCTCTGCGACCGACGAGCTGGGCGCGACCTGCTCTCGGCTGATCCAGCGC 271  
QY 498 GCCAACCTGGGCAACGTGCTCATCTGCCAGCTGACAGCGGAGTCAACAGCGCTTC 557  
DB 270 GCCCGCTGTGCTGTGATCGCGCTGTCTGTGATGATTCGCTATCCCGGCAATC 211  
QY 558 CCCAGGCCCGACGAC 572  
DB 210 CTCCTGGGCTGCTC 196

## RESULT 5

US-09-252-991A-9937  
; Sequence 9937, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

```

: PRIOR APPLICATION NUMBER: US 60/074,768
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 9937
: LENGTH: 1116
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: US-03-252-991A-9937

```

Query Match	2.0%;	Score 45.4;	DB 4;	Length 1116;
Best Local Similarity	48.6%;	Pred. No. 0.035;		
Matches 124;	Conservative	0;	Mismatches 131;	Indels 0;
				Gaps 0;

Qy	318	TACATCAGCTCATATCCCAATTCCTTCGTCGTGGGCTCTACGAGAGCTGTGTGTAAC	377
Db	263	TTATATGCTGATATCGTATTCGTGGGCGTTCCTGCCCCCTGGATGCCCCCTTACATATCCG	322
Qy	378	CGCTGTGGAAACCATGACGAACCTGCGGAGCCGACCGCTCATGAGCGCTGTGTG	437
Db	323	AGGAGAGATTTCCGGGACCTTCTGTGTGATCCCGGCATCTTGATGGAAGCGGCGCAAGCGG	382
Qy	438	GGCTTGTCGAAAGGCAAGAGACGAGAGGCGCGCTGTGCGGCGAGCTATACCGTAC	497
Db	383	CGCTTCTGTGTCGGACCCAGACGATGTGGGCGCCGACCTGTCTTCGCGGTATATCAAGGC	442
Qy	498	GCCAACTGTGGCAACGTGTCTCATCTCTGCGACAGCTACACACGCAAGTCTAACACGCTTC	557
Db	443	GCCCGCCTGTGCTGTGTGATCGGCGCTGTCTTGATGTGTGATCTTGCTATATCCCGGATTC	502
Qy	558	CCGACGGCCCAAGAC	572
Db	503	CTCTCGGCTGCTC	517

```

RESULT 6
US-09-252-991A-9758
: Sequence 9758, Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 31142
: SEQ ID NO 9758
: LENGTH: 1926
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-9758

```

Query Match	2.0%	Score 45.4	DB 4	Length 1925
Best Local Similarity	48.6%	Pred. No. 0.049		
Matches 124	Conservative	0	Mismatches 133	Indels 0
			Gaps 0	
QY	318	TACATCAGCTCATCCCATTTCTTGTGCTGAGCTTCTACGAGCGTGGTGTAC	377	
Db	1129	TTCAATGCTGCTGATCGATGTTCTGGCGCGCTGTTCGCCCCCTGGGTGCGCCCTTACGATCCG	1188	
QY	378	CGCTGCTGGAACCACTATCAGAGAACTGCGCTGGGCCCGACCGCTTCAIAGCTTGGTTCG	437	
Db	1189	AGCAGCAGATTCGCGGACCTTCTCTGTACCCCGCATCTGGCTGGAAGCGGCGACGACGCG	1248	
QY	438	GGCTTCGTCGAAGGCAAGACGACAGGCGCGGCTGCTGGCGGCGCAATCTATCCGCTAC	497	
Db	1249	CGCTTCTCTGCTCGGACACGACGAGCTGGGCGCGCACTGCTTCTGCGGCTGATCCACGCG	1308	

Oy	498	GCCACCTGGGCAACGCTCATCTGCGAGCGTCAGCAACCGAGCTCAACAGCGCTTC	557
Db	1309	GCCCGCTGTACGCGCTGATTCGCGCTCTCTCGGTGTGTATCTCGTGAATCCCCGGCATC	1368
Oy	558	CCGACGCGCCGACGAC	572
Db	1369	CTCTTCGAGCGCTGCTC	1383

RESULT 7  
US-09-252-991A-7100/c  
; Sequence 7100, Application US/09252991A

```

: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,786
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 7100
: LENGTH: 2823
: TYPE: DNA
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-7100

```

	Query Match	1.9%;	Score 43.2;	DB 4;	Length 283;
	Best Local Similarity	52.8%;	Pred. No. 0.24;		
	Matches 93;	Conservative 0;	Mismatches 83;	Indels 0;	Gaps 0;
QY	414	GACCGCCTCATGAGCTGTGTGCGGCTTCTGTGAAGGACGAGACGAGCCGGCTTG	473		
Db	2478	GACCGCTACACGCGTCTGGACGCGCGCGACTGCGCGCGCTGTGGCACTGGCGGACGCG	2419		
QY	474	CTGGGGGCGACGCTCATCCGCTACGCCAACCTGGGCAAGTGTCTATCTCTGGCAACGCTC	533		
Db	2418	TACCAAGTCAACCTTCACAACCTTGGCCACAGCGCGCTGGGCGCTGTGACCTCGCCCGCTTC	2359		
QY	534	AGCACGCAAGTACACAAGGCTTCCGCAAGGCCACACCACTGGTGTGAAGAGGCTT	589		
Db	2358	AGGGGCGAGCGCACAGCTGCTTTCGGCGTACCGTGTGACCGGGGGGTCCGGGTGGGCAT	2303		

```

RESULT 8
US-09-252-991A-7025
; Sequence 7025, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7025
; LENGTH: 3456
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-7025

```

Query Match	1.94;	Score 43.2;	DB 4;	Length 3456;
Best Local Similarity	52.84;	Pred. No. 0.27;		

RESULT 9  
US-09-252-991A-6738  
; Sequence 6738, Application US/09252991A  
; Patent No. 6738

Query Match	1.9%;	Score 42.6;	DB 4;	Length 558;
Best Local Similarity	46.7%;	Pred. No. 0.13;		
Matches 135; Conservative	0;	Mismatches 154;		

RESULT 10  
US-09-621-976-1878  
; Sequence 1878, Application US/09621976  
Patent No. 5,533,000

```

1  GENERAL INFORMATION:
2  APPLICANT: Dumas Milne Edwards, J.B.
3  APPLICANT: Jobert, S.
4  APPLICANT: Giordano, J.Y.
5  TITLE OF INVENTION: ESTs and Encoded Human Proteins
6  FILE REFERENCE: GENSET 054PR2
7  CURRENT APPLICATION NUMBER: US/09/621,976

```

Query Match	1.9%	Score 42.6;	DB 4;	Length 787;
Best Local Similarity	67.4%;	Pred. No. 0.16;		
Matches 60;	Conservative 0;	Mismatches 28;	Indels 1;	

RESULT 11  
US-09-252-991A-6782  
; Sequence 6782, Application US/09252991A

Query Match	1.9%;	Score 42.6;	DB 4;	Length 2346;
Best Local Similarity	46.7%;	Pred. No. 0.31;		
Matches 135; Conservative	0;	Mismatches 154;	Indels 0;	Conserved 0;

QY	281	GCTGATGTTTGGAAATGACTCTGTATTATGGCAGCTTAATCAGGTATATCCCATTTTC	340
Db	1407	GCCGCTACCGGAAAAACCGAACAGGCAAGGCTTGGAGGAGCGCCCGCTACTGCCATCGC	1466
QY	341	CTTCGAGCTGAGGCGCTCTACATGACGCTGTGTGTATCCCGTGTGGAAACAGATACAGAA	400
Db	1467	GACCTCTGCCCAAGCTTGACCTGGACCTGGACCTGGACCTGGAACTTGTGGCGGCTGACAGTGAAGA	1526
QY	401	CCTGCCGTGGACCCGACCGGCTCATAGACCTGTGTGTGGGCTTCTGTGAAGGACAGAGA	460
Db	1527	GCATGCCGATCGACGACGCGCAGACCTGAAGCTGGCGGGCGACAGGCGGGGTATATAGCGCTGA	1586
QY	461	GCAGAGCGCGGAGCTGTGGGCGGCGACGCTCATCCCTCTACGCCAACCTGGGCAACGTGTCTCAT	520
Db	1587	CGACATGCGCGGCGGGCTGTATCAACGCGCGCTTCAACCGCCAAAGGCGCAACTGTGATGTACG	1646
QY	521	CCTGGCGAGCGTCAAGCAACCGAGTTTATCAAGCGCTTCCCGACGGCCGAC	569



Db 1647 CCAGACAGAGCGCGTGTGACCGCCACAAACATCGCAACGTGGCG 1695

RESULT 12  
US-09-252-991A-3386  
; Sequence 3386, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3386  
; LENGTH: 630  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3386

Query Match 1.9%; Score 42.4; DB 4; Length 630;  
Best Local Similarity 49.1%; Pred. No. 0.16;  
Matches 112; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 309 TGGCAGACTACATCCAGCTCATCCCATTTCTTGCTGGGCTTCTACGACGCTG 368  
Db 352 TGGAGCTGCGCCACGACGATCGAACGCTCGCGCTGTCTGCGGGCTGTCTTC 411  
QY 369 GTCTGACCCGCTGTGTGAACAGTACGAGAACCTCGCGCCGACCGCTCATGAGC 428  
Db 412 GTGAGCATGTGACCGGCGCCCGCATCTGCGAGTGCGCTVCAACCTGTTCGCC 471  
QY 429 CTGCTGCGGCTTCTGTCGAAGGACGACGACGAGGCGGTGCTGCGGCCAGGCTC 488  
Db 472 ATGTTCACGCGCGGACGAGGAGATCGAAGGCTATCGCAGCAGATCCCGCTGCTG 531  
QY 489 ATCCGCTACGCCAACCTGGGCAAGTGTCTCATCTCGCGACGCTGACG 536  
Db 532 CTGCGGAGAGACTGTGCGCCCGACGAGATGCTGTAGCAGCCGCAATC 579

RESULT 13  
US-09-252-991A-3344  
; Sequence 3344, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 3344  
; LENGTH: 1647  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-3344

Query Match 1.9%; Score 42.4; DB 4; Length 1647;  
Best Local Similarity 49.1%; Pred. No. 0.29;  
Matches 112; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 309 TGGCAGACTACATCCAGCTCATCCCATTTCTTGCTGGGCTTCTACGACGCTG 368  
Db 212 TGGAGCTGCGCCACGACGATCGAACGCTCGCGCTGTCTGCGGGCTGTCTTC 271  
QY 369 GTCTGACCCGCTGTGTGAACAGTACGAGAACCTCGCGCCGACCGCTCATGAGC 428  
Db 272 GTGAGCATGTGACCGGCGCCCGCATCTGCGCAGTGCGGCTVCAACCTGTTCGCC 331  
QY 429 CTGCTGCGGCTTCTGTCGAAGGACGACGACGAGGCGGTGCTGCGGCCAGGCTC 488  
Db 332 ATGTTCACGCGCGGACGAGGAGATCGAAGGCTATCGCAGCAGATCCGCTGCTG 391  
QY 489 ATCCGCTACGCCAACCTGGGCAAGTGTCTCATCTCGCGACGCTGACG 536  
Db 392 CTGCGGAGAGACTGTGCGCCCGACGAGATGCTGTGTAGCAGCCGCAATC 439

RESULT 14  
US-09-796-766-13  
; Sequence 13, Application US/09796766  
; Patent No. 6660850  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Steve  
; APPLICANT: Lightner, Jonathan  
; APPLICANT: Rafalski, Antoni  
; TITLE OF INVENTION: BRITTLE-1 HOMOLOGS  
; FILE REFERENCE: B1157 US CIP  
; CURRENT APPLICATION NUMBER: US/09/796,766  
; PRIOR FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 09/668884  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: PCT/US99/06583  
; PRIOR FILING DATE: 1999-03-22  
; PRIOR APPLICATION NUMBER: 60/079420  
; PRIOR FILING DATE: 1998-03-26  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 13  
; LENGTH: 1089  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-796-766-13

Query Match 1.9%; Score 42.2; DB 4; Length 1089;  
Best Local Similarity 70.9%; Pred. No. 0.25;  
Matches 56; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 2148 CACACCTAGTACTGCGCCCAACTATGATTTAATAATACAATACTCGTTAAAAA 2207  
Db 1011 CAGACTCTAATTTATTTTCAATCAATGTGTTCTTAATAAAAAAAAAAAAAA 1070  
QY 2208 AAAAAAAAAAAAAAAAAA 2226  
Db 1071 AAAAAAAAAAAAAAAAAA 1089

RESULT 15  
US-09-252-991A-14038/c  
; Sequence 14038, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 14038



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 04:26:28 ; Search time 742.548 Seconds  
(without alignments)  
11175.829 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 cagsgagctccaccagccta.....aaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY\_NTC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*

3: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*

4: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*

5: /cgn2\_6/prodata/2/pubpna/US07\_PUBCOMB.seq:\*

6: /cgn2\_6/prodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*

10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*

11: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*

12: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*

13: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*

14: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*

15: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*

16: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*

17: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*

18: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1247.2	56.0	1263	10	US-09-746-783-3
2	587.8	26.4	1717	15	US-10-264-237-259
3	459.2	20.6	2500	15	US-10-108-260A-1299
4	350.4	15.7	1326	13	US-10-071-766-10
5	215.4	9.7	1350	9	US-09-768-826-15
6	213	9.6	620	10	US-09-814-353-19504
7	189.2	8.5	1198	9	US-10-198-846-11070
8	175	7.9	1198	9	US-09-880-107-2174
9	161.4	7.2	578	15	US-10-027-633-134530
10	149.4	6.7	748	15	US-10-027-633-134529
11	135.2	6.1	539	14	US-10-029-386-7583
12	134.2	6.0	235	14	US-10-029-386-21283
13	116	5.2	853	14	US-10-198-846-11346
14	104	4.7	462	14	US-10-198-846-2561
15	103.6	4.6	3186778	15	US-10-027-632-174961

16	102	4.6	615	14	US-10-198-846-7843	Sequence 7843, Ap
17	98	3.0	507	10	US-09-764-872-195	Sequence 195, App
18	66.2	4.4	442	10	US-09-814-353-14343	Sequence 14343, A
19	66	3.0	748	15	US-10-027-632-134529	Sequence 134529,
20	65.4	2.9	260	10	US-09-814-353-1601	Sequence 1601, Ap
21	65.4	2.9	260	10	US-09-814-353-7959	Sequence 7959, Ap
22	65.4	2.9	463	10	US-09-814-353-13563	Sequence 13563, A
23	64	2.9	751	14	US-10-198-846-1862	Sequence 1862, Ap
24	63.2	2.8	298	10	US-09-814-353-806	Sequence 806, App
25	63.2	2.8	298	10	US-09-814-353-7178	Sequence 7178, Ap
26	60	2.7	60	10	US-09-908-975-16358	Sequence 16358, A
27	54.4	2.4	65	10	US-09-908-975-23941	Sequence 23941, A
28	49.2	2.2	1263	13	US-09-746-783-3	Sequence 3, Appl1
29	49.2	2.2	1326	13	US-10-071-766-10	Sequence 10, Appl
30	48.2	2.2	8607	14	US-10-204-708-72	Sequence 72, Appl
31	48	2.2	16918	14	US-10-311-455-1589	Sequence 1589, Ap
32	47.8	2.1	2201	12	US-10-424-599-133253	Sequence 133253,
33	47.8	2.1	6668	14	US-10-311-455-1670	Sequence 1670, Ap
34	47.6	2.1	466	12	US-10-424-599-10863	Sequence 10863, A
35	47.4	2.1	12601	14	US-10-311-455-2179	Sequence 2179, Ap
36	47.2	2.1	891	14	US-10-156-761-6551	Sequence 6951, Ap
37	47.2	2.1	902568	14	US-10-156-761-1	Sequence 1, Appl
38	47	2.1	2076	10	US-09-986-480-116	Sequence 116, Appl
39	46.8	2.1	5349	14	US-10-311-455-313	Sequence 313, Appl
40	46.8	2.1	3673778	14	US-10-312-841-1	Sequence 1, Appl1
41	46.8	2.1	3673778	14	US-10-312-841-2	Sequence 25, Appl
42	46.6	2.1	846	9	US-09-349-015-25	Sequence 21, Appl
43	46.6	2.1	846	14	US-10-219-664-21	Sequence 7, Appl1
44	46.6	2.1	8195	9	US-09-795-693-7	
45	46.6	2.1	8195	14	US-10-156-239-7	

#### ALIGNMENTS

RESULT 1

US-09-746-783-3

Sequence 3, Appl1

Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: Jacobe, Kenneth

McCoy, John M.

LaValle, Edward R.

Racie, Lisa A.

Treacy, Maurice

Spaulding, Vikki

Agostino, Michael J.

Howes, Steven H.

Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1263 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-746-783-3  
  
Query Match 56.0%; Score 1247.2; DB 10; Length 1263;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 978 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGCAACCACTGATTTGTC 1037  
DB 1 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGCAACCACTGATTTGTC 60  
  
QY 1038 GACAGGAATTTGACAGGTGTCCTGTTGAGTGAGTGAACCAAGACCTGCTCGG 1097  
DB 61 GACAGGAATTTGACAGGTGTCCTGTTGAGTGAGTGAACCAAGACCTGCTCGG 120  
  
QY 1098 ATGAGACCGGACATGTACTGGAATTAAGCCGACACAGCCCTTACACAGCTGCTTCC 1157  
DB 121 ATGAGACCGGACATGTACTGGAATTAAGCCGACACAGCCCTTACACAGCTGCTTCC 180  
  
QY 1158 GCCCAGTTCGGTCAGACCTCCTTTATAGGCTCCACCTTCAACATCAGCTGGAACAAAG 1217  
DB 181 GCCCAGTTCGGTCAGACCTCCTTTATAGGCTCCACCTTCAACATCAGCTGGAACAAAG 240  
  
QY 1218 GAGATGAGATTCAGACCCCAATCAGAGAGACAGAGAGATGCTCAAGCTGAGCATTTGGC 1277  
DB 241 GAGATGAGATTCAGACCCCAATCAGAGAGACAGAGAGATGCTCAAGCTGAGCATTTGGC 300  
  
QY 1278 CGCTTCTAGGCTGCAAGTCCCATGATCAACATCTTCCAGAGGCAAACTGAGGACCAA 1337  
DB 301 CGCTTCTAGGCTGCAAGTCCCATGATCAACATCTTCCAGAGGCAAACTGAGGACCAA 360  
  
QY 1338 CTACTGTGGCCCAAGAGGAATCCCTTCTCCAGAGGSCCTGCCCAAAACCAAGAGCA 1397  
DB 361 CTACTGTGGCCCAAGAGGAATCCCTTCTCCAGAGGSCCTGCCCAAAACCAAGAGCA 420  
  
QY 1398 GCCAACAAGAACGTTAAGGAGGCGCAGAGAGACAAAGGCTGGAAGCTTGAAGCTGTGAC 1457  
DB 421 GCCAACAAGAACGTTAAGGAGGCGCAGAGAGACAAAGGCTGGAAGCTTGAAGCTGTGAC 480  
  
QY 1458 GCCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTACTACAGTGCCTCCACAGAGCCGC 1517  
DB 481 GCCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTACTACAGTGCCTCCACAGAGCCGC 540  
  
QY 1518 CTGAGCCCACTCCCATGTTCTTCCCTAGAAACATCAGGCGCTCAAAAGCTTCAAGT 1577  
DB 541 CTGAGCCCACTCCCATGTTCTTCCCTAGAAACATCAGGCGCTCAAAAGCTTCAAGT 600  
  
QY 1578 GTCAACAGGCAATGACACCAAAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAA 1637  
DB 601 GTCAACAGGCAATGACACCAAAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGCAAGAA 660  
  
QY 1638 AGTTTGAATTTGCTTCAGAGAGAGATGGGCTTGAATGAGAGACCCCAAGATTTTCA 1697  
DB 661 AGTTTGAATTTGCTTCAGAGAGAGATGGGCTTGAATGAGAGACCCCAAGATTTTCA 720  
  
QY 1698 GTGAGAGAGAAAACGTGAGATTAACTGACGATATGACAGAGATCCCGAAAATTCAC 1757  
DB 721 GTGAGAGAGAAAACGTGAGATTAACTGACGATATGACAGAGATCCCGAAAATTCAC 780  
  
QY 1758 CTGAAGAAGACTTTGGAACATACCAACCAATACATACATCACTCACTCAAGATCAGT 1817  
DB 781 CTGAAGAAGACTTTGGAACATACCAACCAATACATACATCACTCACTCAAGATCAGT 840  
  
QY 1818 GATCCTTAATTGGGCTTTGGAAGAACAGGAGTGAAGCATTTCTTAACCTGCTTAAAG 1877

DB 841 GATCCTTAATTGGGCTTTGGAAGAACAGGAGTGAAGCATTTCTTAACCTGCTTCTAATG 900  
QY 1878 GATGCTTGGCCAGACAGGCTCTCACTGTGATGATCAGCAGAGAGCACTGATCCAGTCA 1937  
DB 901 GATGCTTGGCCAGACAGGCTCTCACTGTGATGATCAGCAGAGAGCACTGATCCAGTCA 960  
QY 1938 CAGCCATACAGCTGTCCACACTGAGAGACGCTGTCTTAACACAGCTGAAATGAGTTA 1997  
DB 961 CAGCCATACAGCTGTCCACACTGAGAGACGCTGTCTTAACACAGCTGAAATGAGTTA 1020  
QY 1998 GCTTAATAGATAAATCCACAGACTCTGAGCTTAAATGCTTAACTTCAATAAACT 2057  
DB 1021 GCTTAATAGATAAATCCACAGACTCTGAGCTTAAATGCTTAACTTCAATAAACT 1080  
QY 2058 GTGAAGCTAGACTGAACCACTTGAAGAACTTTAATCACTGATTCAGAGTCGGGA 2117  
DB 1081 GTGAAGCTAGACTGAACCACTTGAAGAACTTTAATCACTGATTCAGAGTCGGGA 1140  
QY 2118 ACCCTTAGTCTATCTGAATCCAGACAGCCACACCTTAGTATCTGCCCAACTAATGA 2177  
DB 1141 ACCCTTAGTCTATCTGAATCCAGACAGCCACACCTTAGTATCTGCCCAACTAATGA 1200  
QY 2178 GTTAATTAATCAAAATCTGTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2229  
DB 1201 GTTAATTAATCAAAATCTGTTAAAAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1252  
  
RESULT 2  
US-10-264-237-259  
; Sequence 259, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Biase et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P431P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; PRIOR FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 259  
; LENGTH: 1717  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1072)..(1072)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1596)..(1596)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1634)..(1634)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1657)..(1657)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1702)..(1702)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1712)..(1712)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-264-237-259

Query Match 26.4%; Score 587.8; DB 15; Length 1717;  
Best Local Similarity 98.8%; Pred. No. 1.2e-157;  
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1242 GAGGACGAGAGATGCTCACGCTGGCATTTGGCCGCTTCTAGACCTGCACTGCCAT 1301  
DB 1 GAGGAGAGAGAGATGCTCACGCTGGCATTTGGCCGCTTCTAGACCTGCACTGCCAT 60  
QY 1302 GATCACCATCTCTCCAGGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAAATCC 1361  
DB 61 GATCACCATCTCTCCAGGGGCAAACTCAAGGACCAAACTACTGTGGCCCAAGAGGAAATCC 120  
QY 1362 CTCTCTCAGAGGGGCTGCCCAAAAACACAGGACCAAAAGAGAGAGAGAGAGAGAGAG 1421  
DB 121 CTCTCTCAGAGGGGCTGCCCAAAAACACAGGACCAAAAGAGAGAGAGAGAGAGAGAG 179  
QY 1422 GAAGACACAGAGGCTGGAAAGCTTAAAGCTGTGAGGCTTCAAGTTGGCCCACTGTAT 1481  
DB 180 GAAGACACAGAGGCTGGAAAGCTTAAAGCTGTGAGGCTTCAAGTTGGCCCACTGTAT 239  
QY 1482 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGGCGCCCTGAGCCCACTCCCATGTTCTTC 1541  
DB 240 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGGCGCCCTGAGCCCACTCCCATGTTCTTC 299  
QY 1542 CCCCTAGAAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGACATAGACCAAAAGAC 1601  
DB 300 CCCCTAGAAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGACATAGACCAAAAGAC 359  
QY 1602 AAAAGCTTAAAGCTGTAGTCTGTGGGCGCAAGAAAAGTTTGAATTTGCTTCAAGAGAC 1661  
DB 360 AAAAGCTTAAAGCTGTAGTCTGTGGGCGCAAGAAAAGTTTGAATTTGCTTCAAGAGAC 419  
QY 1662 GATGGGGGCTGTGTGAGAGACCCAGAAAGTATCTCAAGTGTGAGAGAAAAGTGTGAGATT 1721  
DB 420 GATGGGGGCTGTGTGAGAGACCCAGAAAGTATCTCAAGTGTGAGAGAAAAGTGTGAGATT 479  
QY 1722 AACCTGACGGATATGCGAGAGATCCCGGAAATCACCCTCAAGAACCTTTGGAACATCA 1781  
DB 480 AACCTGACGGATATGCGAGAGATCCCGGAAATCACCCTCAAGAACCTTTGGAACATCA 539  
QY 1782 CCAACCAACATACACTACACTCAAAAGATCACATGATCTTATTTGGGCTTTGAAAAC 1841  
DB 540 CCAACCAACATACACTACACTCAAAAGATCACATGATCTTATTTGGGCTTTGAAAAC 599  
QY 1842 AGGGATG 1848  
DB 600 AGGCTTG 606

RESULT 3  
US-10-108-260A-1299  
Sequence 1299, Application US/10108260A  
Publication No. US20040005560A1  
GENERAL INFORMATION:  
APPLICANT: HELIX RESEARCH INSTITUTE  
TITLE OF INVENTION: No. US20040005560A1el full length cDNA  
CURRENT APPLICATION NUMBER: US/10/108,260A  
NUMBER OF SEQ ID NOS: 5458  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1299  
LENGTH: 2500  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-108-260A-1299

Query Match 20.6%; Score 459.2; DB 15; Length 2500;  
Best Local Similarity 63.5%; Pred. No. 1.3e-120;  
Matches 701; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY 105 ATGACCATCATCTACAGAGCCAAAGTGGCTAATGCCCGCTTAGGCTCTTCTCCGCGCTG 164  
DB 105 ATGACCATCATCTACAGAGCCAAAGTGGCTAATGCCCGCTTAGGCTCTTCTCCGCGCTG 164

DB 229 ATGACTGTCACTTACTCCAGTAAGTAGCAAAATGCAACTTTTTTGGATTTTCATAGATTAA 288  
QY 165 CTGCTGTGCGGGGGGAGCAATCTACAGCTGTATATGGAGTTTAACTTCTTCG 224  
DB 289 CTCTCAAGTGGAGAGCACTTACAACTACCTACAGGGAATTTATGTTTGTCT 348  
QY 225 CTCTGTACTATCATCTCCGCTTATTTATAGGCTGGCCCTCAGGAGAACAGAGCTG 284  
DB 349 GTTCTTTATACAGCAATTAAGTTTGTGTACAGATTTTATCTTACAGAGAGTCAAAACGT 408  
QY 285 ATGTTTGAAGAACTGACTGTATTTGCGACAGCTATCATAGCTATCTCCCATTTCTTC 344  
DB 409 TACTTTGAAAATTAATCAATTTACTGTGACAGTATGCTGAACAAATTTCCAGTAACTTT 468  
QY 345 GTGCTGGGCTTCAAGTGAAGCTGTGAGAACCCGCTGGTGAACAGTACAGAACTG 404  
DB 469 GTGCTGGGCTTCAAGTGAAGCTGTGAGAACCCGCTGGTGAACAGTACAGAACTG 528  
QY 405 CCGTGCCCGAGCCGCTCATGAGCTGTGTGGGCTTGTGCGAAGGACAGAGCA 464  
DB 529 CCGTGCCCGAGAGGCTTATGTTCTCATCTCAGAGAGTTCACGGAAGGACAGAGAC 588  
QY 465 GCGCGGCTGCTGCGGCGACGCTCATCGGCTACGCCAACCTGAGCACTGCTCATCTG 524  
DB 589 GCGCGGCTGCTTAAAGAGAGCTGATGCGCTACGTCAATCTCACTCCGTGCTCATCTTT 648  
QY 525 CCGAGGTGAGACCGCACTTCAAGGCGCTTCCCGAGGCGCCAGACCTGTGTGCAAGCA 584  
DB 649 CCGTGGTGAAGACCTGTGTGACAAAGATTTCCCAATGAGACCAAGTGTGTAAGCA 708  
QY 585 GCGTTATGACTCCGCGCAACACAGAGCTTGAAGAACTGAGCTTACCAACAATG 644  
DB 709 GCGTTTATACAAAGATGAAAGAAATTAATTCACACCTCAAGCTCTCATCTGAA 768  
QY 645 TTCTGGGTCCTGGTGTGTGTTGCAACTGTCTAATGAGGCGTGTGAGAGTCA 704  
DB 769 TATTTGGGTTCCATTCATCTGTGTTGAAATCTTGCACCTAAAGCCCGGATGAAAGTAA 828  
QY 705 ATCCGGAACCTATCCTGTCCAGAGCCCTGTGAACGAAATGAACCTTGGTATCTGAG 764  
DB 829 ATCAGAGAGATGTTGATCTGCAATCTTAATGATGATGAATGAAATGCAATCCGCTTGG 888  
QY 765 TGTGACACTGTATGCTACAGACTGATTAATATCCATGCTGTATACAGAGTGTG 824  
DB 889 TGCAGCTTATATGCTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 948  
QY 825 ACTGTGGCGGTGACACTTCTTCTGACTGTGATGATGATGATGATGATGATGATGATG 884  
DB 949 ACTGTGTGTGTATATCTTCTTCTTGTGCTGCTGATGATGATGATGATGATGATGATG 1008  
QY 885 GCCAAGGCTTACCTGTGAGAGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 944  
DB 1009 ACCAAGGCTTACGAGGCGATATGATGATGATGATGATGATGATGATGATGATGATG 1068  
QY 945 TTCTTCTTATGTTGGCTGTGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1004  
DB 1069 TTCTTCTTATATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1128  
QY 1005 GATGATGATGATTTTGAACCAACTGATTTGTGACAGGAAATTTGAGAGTGTCTCTG 1064  
DB 1129 GATGATGATGATTTTGAACCAACTGATTTGTGACAGGAAATTTGAGAGTGTCTCTT 1188  
QY 1065 GCTGTGATGATGATGACAGAGAGCTGCTGCGATGAGAGCGGACATGATGATGATG 1124  
DB 1189 GCTGTGAGAGAAATGACATGAGCTTACCAAGATGAAAGAGACATTTTACTGAGAG 1248  
QY 1125 CCGAGGCGACAGCCCGCTTACAGAGCTGCTGCGCCAGTTCGCTGAGAGCTCTTATG 1184  
DB 1249 TCTGCTGCTGCGCCACATACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1308  
QY 1185 GCGTCCACCTTCAACATGAGCTG 1208  
DB 1309 GCGTCCACATGATGAGAGCTG 1332



```
Db 143 CTCGTCTACTTTGGAGAGCTGAACAGTACCGAGCCAGTGCAGCATGTCTATTCCAC 202
Qy 783 TACAGCTGATAGTATCCCACTGATGTATACAGAGTGTGACTGNGCGGTGTACAG 842
Db 203 TATGACTGATCAGCATCCCTCTGCTTACACCCAGGTGTGATGACATGCGTCTACTCT 262
Qy 843 TTCTTCTGACTGTCTAGTGTGGCGGCACTTTCTGAAACCAAG-----885
Db 263 TTCTTTGCCCTCTCCCTGTTGGCCGCACTTTGTGAGCAGAGG-AGGGGCTGCGAAA 322
Qy 886 -----CCAAAGCCTTACCTTGCCCATGAG---CTGAACTTC 917
Db 323 CCTCAGAGAGCTTCTGAAGCCAGGCGCAGAGCCAGCCCTGGAGAGACCCGACATG 382
Qy 918 GTTGTGCGCGTCTTCAAGCTTCTGCAAGTCTTCTGTATGTGGTGGTGGTGGAGTGGCA 977
Db 383 TACGTGCTTCAACACTGTCTGCAAGTCTTCTGTATGTGGTGGTGGTGGAGTGGCT 442
Qy 978 GAGCAGCTCATCAACCCCTTGGAGAGATGATGATGATTTTGAACCAACTGGAATTGTC 1037
Db 443 GAAACATATCATCAACCATTTGGTGGAGATGATGAGACTTTGAGACAAATCATGCTCAT 502
Qy 1038 GACAGAGATTTGAGAGTGTCTCTGTGTGGCTGTGATGATGATGCAACGAGACTGCTCG 1097
Db 503 GACCGCAACTTGAGAGGTGTCTCTGTATCCGTGAGAGAAATGTACGAACTTCCCGCC 562
Qy 1098 ATGAGACCCGAGCATGTATGGAATTAAGCCGAGCCAGCCCGCCCTTACAC--AGCTGCT 1154
Db 563 GCTGAAGAGAGACAGTACTGGATGAGAGACAGCCGAGACCATCTGATGTGGACAG 622
Qy 1155 TCCGCCAGTTCCTGAGAGCTCTTATATGGGCTTCCACCTTCAACATCAGACCTGAAACA 1214
Db 623 GCGGCGGAGTCTGTGGCGCCCTCATTCCTGTGGGCTTCAACCTTGCGGATGAGCGAC 682
Qy 1215 GAGGAGATGAGATTCCAGC 1233
Db 683 GACCTGAGCAGAGCCTGCG 701

RESULT 6
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814, 353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504
Query Match 9.6%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 2,9e-50;
Matches 306; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 103 CCATGACCATCACTTACAGAGCCAGGCTTAATGCCCCGCTTAGGCTCTCCCGCC 162
Db 464 CCATGACGGTTCATCACTTCAAGTGGAGGCGCGCTTCGAGAGTTCCTGCGCC 405
Qy 163 TGTGCTGTGTGAGCGGCGAGCATCTCAAGCTGTATATGCGAGTCTTATCTTCC 222
Db 404 TGTCTTCCGCTGAGAGGGAAGATCTCAAGTCTCTTCAAGAAATCTCTCTTTG 345
Qy 223 TGTCTGTCTAATCATATCCGCTTATTTATATATGCTGCGCCTCAGCGAAGAACAGC 282
Db 344 GGGCTTGTATACGCTGTGTATGATCACCACCGGCTGCTGTGACCCAGAGCAGAGGT 285
Qy 283 TGAATTTGAGAACTGACTGTATTTGAGACAGTCAATCCAGCTATCCCATTTCT 342
Db 284 AGTGTATCTAGTGGTGGCGGCTGATGCAACCGCTCAGACAGCTCATTTCTTCT 225
Qy 343 TGTGCTGAGGCTTCTAAGTGAAGCTGTGTGACCCGCTGTGTGAACAGTACAGAGAC 402
Db 224 TGTATTTGGGTTTCTATATGATCTCTGTGTGAACCGCTGTGTGTCCAGTACAGAG 165
Qy 403 TGTGCTGAGCGGAGCCGCTCATAGAGCTGTGTGTGGCTTGTGAGAGCAAGAGAGC 462
Db 164 TCCCGGTGACAGACAGCTGATGTGTGCTATCTCGGCTAGCGGACCGCTGTGAGAC 105
Qy 463 AAGCGCGGCTGTGTGGGCGGACGCTCATCCGCTTACCGCAACCTGGGCAAGCTCAT 522
Db 104 GGGCGCGCTGTGTGGCGGCGGCTCATCTGCTACCGCAAGCTGTGTGTGTGTGTG 45
Qy 523 TGTGAGCGTCAAGACCGGCTTACAGCGCTTCCCGAGC 563
Db 44 TGTGCTGTGTGAGCAGCGCGGCTGTATAGCGCTTCCCGAC 4

RESULT 7
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1287_1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 8.5%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 3.1e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
```

OY	349	GGGCTTCTACGTGAGCGTGGTCTCGGACCCGCTGGGTGGAAACCAAGTACGAGAACCTGGCGT	408
Db	318	TAGGGTTTATATGTACTCTGGTAGTGAACCGATGGTGGAAACGATTTGTGAATTTCCCT	377
OY	409	GGCCGACGCGCTCATAGAGCTGTGTCCGGGCTTCGTGGAAGGCAAGGACGAGCAAGGCC	468
Db	378	GGCCAGACAGGCTTAATGTTCCCTCATCTTAGCAGTGTTCACGGAAACGAGACAGCGGG	437
OY	469	GGCTGCTGGCGGCAAGCTCATTCGGTACGCCAACCTGGGCAACGTGTCATTCCTGGCA	528
Db	438	GCTGCTTATAGAAGAGACGCTGATGCGGTACGTCAATGTCACTCCGCTCATTTTCGCT	497
OY	529	GCGTCAGCACGCGAGCTCTACAGCGCTTCCCCAGCGCCAGCACCTGATGCAAGCAGGCT	588
Db	498	CGGTAGCAGCTCTGTGTTCAAAAGTTTCCCAATGGAACACAGTGGTTGAAGCAGGTT	557
OY	589	TTATGACTCCGGCAGAAACAAGCAGTGGAGAACTGAACCTTACCAACACATGTTCT	648
Db	558	TTATGACAAACAGATGAAGGAATTTTTCACACACCTCAAGTCCCTCATCTGAAATATT	617
OY	649	GGGTGCCCTGGGTGTGTGTTCACACTGTCAATGAAAGCGTGGCTTGGAGGTGCATCC	708
Db	618	GGGTTCACATTACTGTGTGTGAAACTTTCGCACTTAAAGCCGGAATGANGTGAATCA	677
OY	709	GGGACCTTATCCTGTCTCCAGAGCCTGCTGACAGAGATGAAACCTTGCTACTCAGTGTG	768
Db	678	GAGACAGGTGTATCTGCGAATCATTTATGATGCTGAATGAAATGCATACCGCTTTGGTGCA	737
OY	769	GACACCTGTATGCCCTTACAGACTGGATTATGATCCCACTGGTGTATACACAGGTGG	822
Db	738	GCGCTTATTCGTTATGACTGGGTGGGATTCGCTGGTTTACACCCAGAGTGG	791

```

RESULT 8
US-09-880-107-2174/C
: Sequence 2174, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2174
: LENGTH: 1198
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
US-09-880-107-2174
US-09-880-107-2174
US20020142981A1 L20941

```

	Query Match	7.9%	Score 175	DB 9	Length 1198
	Best Local Similarity	100.0%	Pred. No. 3.6e-39		
	Matches 175	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	2027	CAGCCTTTAAAGCCTTTTATTATCAAAAAAATCTGTAAAGCTAGACTGAACCATTTGAAACA			20866
DB	1198	CAGCCTTTAAAGCCTTTTATTATCAAAAAAATCTGTAAAGCTAGACTGAACCATTTGAAACA			11399
QY	2087	TTTAACTCAGACTCTCGATTCAGAGATCGGGAAACCTTGGTCTATCTGAATCCAAAGACAG			21466
DB	1138	TTTAACTCAGACTCTCGATTCAGAGATCGGGAAACCTTGGTCTATCTGAATCCAAAGACAG			10799

[illegible]

```

RESULT 9
US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218
PRIORITY FILING DATE: 2000-02-24
PRIORITY APPLICATION NUMBER: US 60/167,363
PRIORITY FILING DATE: 1999-11-22
PRIORITY APPLICATION NUMBER: US 60/156,358
PRIORITY FILING DATE: 1999-09-28
PRIORITY APPLICATION NUMBER: US 60/146,002
PRIORITY FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

```

	Query Match	Best Local Similarity	7.2%	Score 161.4;	DB 15;	Length 578;
	Matches	162;	Conservative	0;	Fried. No. 1.8e-35;	Indels 0; Gaps 0
QY	582	GCAGGCTTATGACCTCCGGCAGAAACACAGAGAGTTGGAGAAACTAGCCTTACACACAAAC				641
DB	183	GCAGGCTTATGACTCCGGCAGAAACACAGAGAGTTGGAGAAACTAGCCTTACACACAAAC				124
QY	642	ATGTTCTGGAGTCCCTCGGAGTGAGTTTCCCAACCTGTCAATGAAGCGCTTGAGAGGT				701
DB	123	ATGTTCTGGAGTCCCTCGGAGTGAGTTTCCCAACCTGTCAATGAAGCGCGTCTTGAGAGGT				64
QY	702	CGAATCCGGAGCCTTATCTCTGCTCCAGAGCCTGCTGGAACGAGA				744
DB	63	CGAATCCGGAGCCTTATCTCTGCTCCAGAGCCTGCTGGAACGCTTA				21

```

RESULT 10
US-10-027-632-134529/c
Sequence 134529, Application US/10027632
Publication NO. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218

```



```

: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: PatSeq for Windows Version 4.0
: SEQ ID NO 134529
: LENGTH: 748
: TYPE: DNA
: ORGANISM: Human
US-10-027-632-134529

```

Query Match	6.7%	Score 149.4	DB 15	Length 748
Best Local Similarity	90.9%	Pred. No. 6.1e-32		
Matches 170	Conservative 0	Mismatches 16	Indels 1	Gaps 1

Qy	558	CCGAGGCCCGCAGCACTGGGTGCAAGAGGCTTTATGATCTCCGGCGAGAACAAACAGTTG	61
Db	712	CCCAACGGCTTTTCACTCCACTCTCGAGGCTTTATGATCTCCGGCGAGAACAAACAGCAG-TG	65
Qy	618	GAGAAACTGAGCCTTACCAACAACATGTTCTGGGTGCCTTGGTGTGGTTTGCACCTG	67
Db	653	GAGAACTGAGCCTTACCAACAACATGTTCTGGGTGCCTTGGTGTGGTTTGCACCTG	59
Qy	678	TCATGAAAGCGTGGCTTGGAGTGCATATCCGGACCTTATCTGTCTCCAAGCTGTGTG	73
Db	593	TCATGAAAGCGTGGCTTGGAGTGCATATCCGGACCTTATCTGTCTCCAAGCTGTGTG	53
Qy	738	AACGAG	74
Db	533	AACGTGA	527

```

RESULT 11
US-10-029-386-7583/C
: Sequence 7583, Application US/10029386
: Publication No. US20030194704A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
: TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
: FILE REFERENCE: AEMICA-X-2
: CURRENT APPLICATION NUMBER: US/10/029,386
: CURRENT FILING DATE: 2001-12-20
: NUMBER OF SEQ ID NOS: 34288
: SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
: SEQ ID NO 7583
: LENGTH: 539
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO CHR19.3
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
: OTHER INFORMATION: EST HUMAN HIT: A4573517.1, EVALUE 9.00e-56
: OTHER INFORMATION: NT HIT: g115304094, EVALUE 1.00e-130
: OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUE 2.00e-25
: US-10-029-386-7583

```

Query Match	6.1%	Score 135.2	DB 14	Length 559
Best Local Similarity	73.3%	Pred. 15.9e-28		
Matches	173	Conservative	0	Mismatches 63
			Indels	0
			Gaps	0
Oy	351	GGCTTCTACGTGACGCTGTCGTGACCGCGTGTGTGAAACAGTACAGAACCTGCGCTGG	410	
Ob	362	GGCTTTATGTGACGCTGTGTGTGGAACCGCTGTGTGTGAAACAGTACATATGTGATGTCGGCTGG	303	

QY	411	302	471	242	QY	531	182
	CCGCAACGGCTATAGACCTGGTGTGCGGGCTTTCGCGAAGGCAGAGACGAGCAAGGCGG	CCGCAACGGCTATAGCTGTGTGTGGCGGGCACCGTGCAACGACCGACGACCGGGGCGG	CTGCTCTGGCGGACGCTATCCGCTACCGCCAAACCTGGGCAACGTCATCTCTGGCAGC	CTCTACCGGCGGACACTCATGTGGCTTACGCAAGGGCTCTGGGCGCGTCTCATCTTGGCTCC		GTCAGCACCGCAGTCTTACAGGCGCTTCCCGACGGCCGACGACCTGTGTCAAGCAGG	GTCAGCACCGGCGGTTCAGAGCGCTTCCCGACCATATAGACACGTGTGTGAGAGCGTGG
	470	243	530	183		586	127

```

RESULT 12
US-10-029-386-21283/c
Sequence 21283, Application US/10029386
Publication No. US20030194704v1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21283
LENGTH: 235
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR19.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 4.00e-56
OTHER INFORMATION: NT HIT: g18923136, EVALUE 1.00e-129
OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUE 1.00e-24
US-10-029-386-21283

```

```

Query Match      6.0%; Score 134.2; DB 14; Length 235;
Best Local Similarity 73.2%; Pred. No. 6,9e-28;
Matches 172; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

OY 352 GCCTCTACGAGCGCTGCTGTGACCCGGCTGTGTGAACCACTAGAGAACTGCGTGGC 411
    |||||
Db 235 GCTTTTATGTGAGCGCTGTGTGTGAACCGCTGTGTGAGACCACTTATGCAATGCGCGTGC 176
    |||||

OY 412 CCGACCGGCTCATGAGCGCTGTGTGCGGGCTTGGTGAAGGCAAGACGAGCAAGCGCGCG 471
    |||||
Db 175 CCGACCGGCTCATGTGTGTGTGTGTGGGACCGGTGACGAGCGAGCGACCGCGCGCGC 116
    |||||

OY 472 TGCTGCGGCGGACGCGCTCATCGGCTACGCACTGTGGGCAACGTGTCTATCTTGGCGACG 531
    |||||
Db 115 TCTACCGGCGGACACTATGCGGCTTACGAGGGCTTCTGGGCCGTCTCATCTTGGGCTTCG 56
    |||||

OY 532 TCAGCACCGGACGCTTACAGACGCTTCCCGACGCGCCGACCTGTGTGACAGCAG 586
    |||||
Db 55 TCAGCACCGGCGGTTCAGACGCTTCCCGACCACTAGACCACTGTGTGAGGCTGG 1
    |||||

RESULT 13
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen

```

APPLICANT: Steinmann, Kathleen  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
FILE REFERENCE: MRI-049  
CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11346  
LENGTH: 853  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,  
LOCATION: 43, 50, 51, 54, 57, 853  
OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-11346

Query Match 5.2%; Score 116; DB 14; Length 853;  
Best Local Similarity 68.2%; Pred. No. 2.6e-22;  
Matches 161; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 350 GGGCTTACGTACGCTGCTGACCCGCTGCTGGAACCACTGACGAACCTGCGT 409  
DB 272 GGGGTTTATGTACTCTGTGTAAGCACTGCTGGAACCACTGTAATTTGCGCTG 331  
QY 410 GCCGACCGGCTCATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469  
DB 332 GCCAGACAGGCTATGTTCTCTCATCTAGCAAGTTCAGGAGCAACAGCAAGGCG 331  
QY 470 GCTGCTGCGGCGACGCTCATCGCTACGCAACCTGCGCAACGCTGCTCATCTG 529  
DB 392 CTTGCTTAAAGAGACGCTGATGCGCTACATCTCACCCTGCTCATCTTTGCTC 451  
QY 530 CTTGACAGACCGGCTGCTACAGGCTTCCCGAGGCCAGCAGCTGCTGCAAGCAG 585  
DB 452 GGTAGACACTGCTGTGTACAAAGATTTCACCAATGACACAGTGTGTAAGCAG 507

## RESULT 14

US-10-198-846-2561/C  
Sequence 2561, Application US/10198846  
Publication No. US2003009974A1  
GENERAL INFORMATION:  
APPLICANT: Lillie, James  
APPLICANT: Xu, Yongyao  
APPLICANT: Wang, Youzhen  
APPLICANT: Steinmann, Kathleen  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
FILE REFERENCE: MRI-049  
CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2561  
LENGTH: 462  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 390, 462  
OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-2561

Query Match 4.7%; Score 104; DB 14; Length 462;

Best Local Similarity 67.6%; Pred. No. 4.9e-19;  
Matches 146; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 349 TGGGCTTACGTACGCTGCTGCTGACCCGCTGCTGGAACCACTGACGAACCTGCGT 408  
DB 255 TAGGTTTATGTACTCTGTGTAAGCACTGCTGGAACCACTGTAATTTGCGCT 196  
QY 409 GGGCGACCGCTCATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468  
DB 195 GGCACACAGGCTAATGTTCTCATCTTACAGAGTGTTCAGAGACGACGAGGCG 136  
QY 469 GCGTCTGCGGCGACGCTCATCGCTACGCAACCTGCGCAACGCTGCTCATCTGCGCA 528  
DB 135 GCGTCTTAAAGAGACGCTGATGCGCTACGTCAATCTCACTCCCTGCTCATCTTCCGT 76  
QY 529 GCGTACGACCGGACTCTTACAGCGCTTCCACGCG 564  
DB 75 CCGTAGACACTGCTGTGTACTCTGCGCGCACCGCG 40

## RESULT 15

US-10-027-632-174961/C  
Sequence 174961, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMOPLHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 174961  
LENGTH: 3186778  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (1)...(3186778)  
OTHER INFORMATION: n = A,T,C or G  
US-10-027-632-174961

Query Match 4.6%; Score 103.6; DB 15; Length 3186778;  
Best Local Similarity 77.5%; Pred. No. 1.5e-16;  
Matches 138; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

QY 418 GCGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477  
DB 285842 GCGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285784  
QY 478 GCGGACGCTCATCGCTACGCAACCTGCGCAACGCTCATCTCTGCGACGCTGCA 537  
DB 285783 GAGCATGTCTCTCGCTTCCGCAACGCGCAACGCTCATCTCTGCGACGCTGCA 285724  
QY 538 CCGCAGTCTACAGGCTTCCCGACGCGCGCAACGCTGCTGCAACGCTTATGAC 595  
DB 285723 CCGTACACAGGCTTCCCGACGCGCGCAACGCTGCTGCAACGCTTATGAC 285666

Mon Mar 29 09:56:48 2004

us-09-622-964-2.rnpb

Page 9

Search completed: March 26, 2004, 11:54:34  
Job time : 760.881 secs

---

This page blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 25, 2004, 22:13:32 ; Search time 170.702 Seconds  
(without alignments)  
7896.639 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429

Sequence: 1 caggagagccacacagcccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: \*  
2: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfile1.seq: \*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.6	3.5	283	4 US-08-579-445-26	Sequence 26, Appl
2	84.6	3.5	5590	3 US-09-050-159-129	Sequence 129, App
3	84.6	3.5	9365	4 US-09-608-285A-8	Sequence 8, Appl
4	84.6	3.5	9365	4 US-09-350-836B-8	Sequence 8, Appl
5	84.6	3.5	9365	4 US-09-370-265-8	Sequence 8, Appl
6	84.6	3.5	9365	4 US-09-557-800C-8	Sequence 8, Appl
7	84.6	3.5	9365	4 US-09-370-625A-8	Sequence 8, Appl
8	84.6	3.5	14747	4 US-09-608-285A-42	Sequence 42, Appl
9	84.6	3.5	14747	4 US-09-557-800C-42	Sequence 42, Appl
10	84.6	3.5	15977	4 US-09-608-285A-59	Sequence 59, Appl
11	84.6	3.5	21784	4 US-09-820-002-3	Sequence 3, Appl
12	84.6	3.5	50000	4 US-09-146-053-3	Sequence 3, Appl
13	84.6	3.5	11531	1 US-08-068-945A-1	Sequence 1, Appl
14	84.6	3.5	11531	1 US-08-442-806-1	Sequence 1, Appl
15	84.6	3.5	11531	4 US-09-355-295B-1	Sequence 1, Appl
16	84.6	3.5	11531	4 US-09-355-295B-1	Sequence 1, Appl
17	84.6	3.5	74962	4 US-09-685-853A-3	Sequence 3, Appl
18	84.6	3.5	786431	4 US-09-751-389-3	Sequence 3, Appl
19	83.2	3.4	36741	3 US-09-301-665-3	Sequence 3, Appl
20	83.2	3.4	41100	4 US-09-755-665-46	Sequence 46, Appl
21	83.2	3.4	319	4 US-09-621-976-13089	Sequence 13089, A
22	83.2	3.4	331	4 US-09-621-976-12648	Sequence 12648, A
23	83.2	3.4	1701	3 US-09-078-294-9	Sequence 7, Appl
24	83.2	3.4	2115	1 US-08-395-800A-7	Sequence 7, Appl
25	83.2	3.4	6769	1 US-08-480-784-20	Sequence 20, Appl
26	83.2	3.4	6769	1 US-08-483-553-20	Sequence 20, Appl
27	83.2	3.4	6769	1 US-08-487-002-20	Sequence 20, Appl
				US-08-483-554B-20	Sequence 20, Appl

28	83	3.4	6769	1 US-08-488-011B-20	Sequence 20, Appl
29	83	3.4	6769	3 US-08-850-727-20	Sequence 20, Appl
30	83	3.4	6769	5 PCT-US95-10202-20	Sequence 20, Appl
31	83	3.4	6769	5 PCT-US95-10203-20	Sequence 20, Appl
32	83	3.4	6769	5 PCT-US95-10220-20	Sequence 20, Appl
33	83	3.4	14636	3 US-09-173-914-6	Sequence 6, Appl
34	83	3.4	16891	4 US-09-486-147-1	Sequence 1, Appl
35	83	3.4	20674	4 US-09-641-638-651	Sequence 651, App
36	83	3.4	70000	4 US-09-851-896-3	Sequence 3, Appl
37	83	3.4	112132	4 US-09-741-150-3	Sequence 3, Appl
38	83	3.4	112132	4 US-10-160-187-3	Sequence 3, Appl
39	82.6	3.4	282	1 US-08-133-629-8	Sequence 8, Appl
40	82.6	3.4	1613	2 US-08-812-204-1	Sequence 1, Appl
41	82.6	3.4	8220	4 US-09-797-908-3	Sequence 3, Appl
42	82.6	3.4	11725	2 US-08-756-506-1	Sequence 1, Appl
43	82.6	3.4	11725	4 US-09-328-925-50	Sequence 50, Appl
44	82.6	3.4	18853	4 US-09-820-005-3	Sequence 3, Appl
45	82.6	3.4	38564	4 US-09-734-673-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-08-579-445-26  
; Sequence 26, Application US/08579445  
; Patent No. 6566053  
; GENERAL INFORMATION:  
; APPLICANT: Perucho, Manuel  
; APPLICANT: Peinado, Miguel A.  
; APPLICANT: Ionov, Yuri  
; APPLICANT: Malkhosyan, Sergei  
; TITLE OF INVENTION: Identification of Neoplasms by Detection  
; TITLE OF INVENTION: Of Genetic Deletions  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Knobbe, Martens, Olson & Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Releasee #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/579,445  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/152,484  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kirkpatrick, Anita M.  
; REGISTRATION NUMBER: 32,617  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 235-8550  
; TELEFAX: (619) 235-0176  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 283 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-579-445-26  
Query Match 3.5%; Score 84.6; DB 4; Length 283;

```
Best Local Similarity 90.9%; Pred. No. 4,7e-12;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 129 AAAAATTAGCGCGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 188
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 1177
    |||||
DB 189 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 227
    |||||

RESULT 2
US-09-050-159-129/c
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Leif T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; EARLIER FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; EARLIER FILING DATE: 1987-04-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 5590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
US-09-050-159-129

Query Match
Best Local Similarity 90.9%; Score 84.6; DB 3; Length 5590;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 1926 AAAAATTAGCGCGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1867
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 1177
    |||||
DB 1866 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 1828
    |||||

RESULT 3
US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
```

```
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match
Best Local Similarity 84.8%; Score 84.6; DB 4; Length 9365;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 7598 AAAAATTAGCGCGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 7539
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 1177
    |||||
DB 7538 AGGAGATGCGCTTGAACCCGGAGGCGGAGGCTTGTGGTG 7500
    |||||

RESULT 4
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; CURRENT FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
```



```

FILE REFERENCE: 28110/36457
CURRENT APPLICATION NUMBER: US/09/557,800C
CURRENT FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/481,238
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: 09/370,265
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273447
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 09/122449
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 09/244444
PRIOR FILING DATE: 1999-02-04
PRIOR APPLICATION NUMBER: 09/118,205
PRIOR FILING DATE: 1998-07-16
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 9365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (3405)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
FEATURE:
NAME/KEY: misc feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-557-800C-8
Query Match 3.5%, Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 3e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
QY 1079 AGAAGCAGCCAGGAGTGTCGCGACACCTGTATCCCACTGCTGGGAGGCTGAGGC 1138
DB 7598 AAAAATTTAGCCRGCGGTGTGCGCGCTGTATCCCACTGCTGGGAGGCTGAGGC 7539
QY 1139 AGAGAAATCGCTTGAACCCCGGAGGCGGAGGTTGTGTG 1177
DB 7538 AGRAGAAATCGCTTGAACCCAGAGAGGAGGTTGTGAGTG 7500
RESULT 7
US-09-370-625A-8/c
Sequence 8, Application US/09370625A
Patent No. 6600032
GENERAL INFORMATION:
APPLICANT: Ford, John
APPLICANT: Mulero, Julio
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
FILE REFERENCE: 28110/35908
CURRENT APPLICATION NUMBER: US/09/370,625A
CURRENT FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: PCT/US99/16180
PRIOR FILING DATE: 1999-07-16
PRIOR APPLICATION NUMBER: 09/350,836
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: 09/273,447
PRIOR FILING DATE: 1999-03-19

```

```

: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 8
: LENGTH: 9365
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1)..(288)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1281)..(1580)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1820)..(1855)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (2467)..(2555)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (2863)..(2942)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (3889)..(3950)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (4894)..(4995)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (5847)..(5987)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (6966)..(7138)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (8556)..(9365)
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (3409)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9214)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9303)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9311)
: OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match      3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 3e-11; 7; Indels 0; Gaps 0;
Matches 84; Conservative 8; Mismatches

QY      1079 AGAAGAGCGCAGCGCTGTGTGCGCACACCTGTATATCCAGTACTCGGAGGCTGAGGC 1138
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       7598 AAAAATTATGCGRGGCGTGTGTGGCGCRYGCCTGTATATCCAGTACTTGGAGGCTGAGGC 7539
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      1139 AGGAGATTCGCTTGAATCCCGGAGGCGGAGGCTTGTGTG 1177
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       7538 AGRAGATTCGCTTGAAACCCAGAGGAGGTGACGTG 7500
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-09-608-285A-42/c
: Sequence 42, Application US/09608285A
: Patent No. 6335013
: GENERAL INFORMATION:
: APPLICANT: Ford, John

```



```
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
/ TITLE OF INVENTION: POLYPEPTIDES
/ FILE REFERENCE: 28110/36570
/ CURRENT APPLICATION NUMBER: US/09/608,285A
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/583,231
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 09/557,800
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350,836
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 09/273,447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/244,444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/122,449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 42
/ LENGTH: 14747
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (13641)
/ OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

Query Match          3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 3.8e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGAGCGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10958 AAAAATTAGCCGCGCGCTGTGGCGCCTGTAAATCCAGCTACTCGGAGGCTGAGGC 10899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1139 AGAGAAATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10898 AGRAGAATCGCTTGAACCCAGGAGGTYGRAGGTGCAGTG 10860

RESULT 9
US-09-557-800C-42/c
/ Sequence 42, Application US/09557800C
/ Patent No. 6476211
/ GENERAL INFORMATION:
/ APPLICANT: Ford, John
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 28110/36457
/ CURRENT APPLICATION NUMBER: US/09/557,800C
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350836
/ PRIOR FILING DATE: 1999-07-09
```

```
/ PRIOR APPLICATION NUMBER: 09/273447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/122449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/244444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 42
/ LENGTH: 14747
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (13641)
/ OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-557-800C-42

Query Match          3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 3.8e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGAGCGGTGGCGCACACCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10958 AAAAATTAGCCGCGCGCTGTGGCGCCTGTAAATCCAGCTACTCGGAGGCTGAGGC 10899
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1139 AGAGAAATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 10898 AGRAGAATCGCTTGAACCCAGGAGGTYGRAGGTGCAGTG 10860

RESULT 10
US-09-608-285A-59/c
/ Sequence 59, Application US/09608285A
/ Patent No. 6335013
/ GENERAL INFORMATION:
/ APPLICANT: Ford, John
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
/ TITLE OF INVENTION: POLYPEPTIDES
/ FILE REFERENCE: 28110/36570
/ CURRENT APPLICATION NUMBER: US/09/608,285A
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/583,231
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 09/557,800
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350,836
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 09/273,447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/244,444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/122,449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 59
/ LENGTH: 15977
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
```

OTHER INFORMATION: CD39-14/L66 Gene Sequence

NAME/KEY: CDS  
LOCATION: (245)..(461)  
NAME/KEY: CDS  
LOCATION: (1454)..(1533)  
NAME/KEY: CDS  
LOCATION: (2734)..(2877)  
NAME/KEY: CDS  
LOCATION: (4364)..(4439)  
NAME/KEY: CDS  
LOCATION: (4679)..(4714)  
NAME/KEY: CDS  
LOCATION: (5326)..(5414)  
NAME/KEY: CDS  
LOCATION: (5723)..(5802)  
NAME/KEY: CDS  
LOCATION: (6751)..(6812)  
NAME/KEY: CDS  
LOCATION: (7758)..(7859)  
NAME/KEY: CDS  
LOCATION: (8712)..(8852)  
NAME/KEY: CDS  
LOCATION: (9831)..(9887)  
NAME/KEY: CDS  
LOCATION: (11613)..(11728)  
NAME/KEY: CDS  
LOCATION: (13146)..(13691)  
NAME/KEY: CDS  
LOCATION: (15702)..(15839)  
NAME/KEY: misc\_feature  
LOCATION: (14871)  
OTHER INFORMATION: n = a or c or g or t  
US-09-608-285A-59

Query Match  
Best Local Similarity 84.8%; Pred. No. 4e-11;  
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGCTGTGGCGCACACCTGTAATCCCGACTACCGGAGGCTGAGGC 1138  
DB 12188 AAAAATTAGCCRCGCGGTGGCGCGCTGTAATCCCGACTACCTGAGGCTGAGGC 12129  
QY 1139 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGGTG 1177  
DB 12128 AGRAGAATCGCTTGAACCCGAGGCGGAGGTTGTGCAGTG 12090

RESULT 11  
US-09-820-002-3/C  
Sequence 3, Application US/09820002  
Patent No. 6482630  
GENERAL INFORMATION:  
APPLICANT: Gan, Weinhu  
APPLICANT: Ye, Jane  
APPLICANT: Difrancesco, Valentina  
APPLICANT: Beasley, Ellen  
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
FILE REFERENCE: CL001194  
CURRENT APPLICATION NUMBER: US/09/820,002  
CURRENT FILING DATE: 2001-03-29  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 21784  
TYPE: DNA  
ORGANISM: HUMAN  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(21784)  
OTHER INFORMATION: n = A,T,C or G  
US-09-820-002-3

Query Match  
Best Local Similarity 83.5%; Pred. No. 4.7e-11;  
Matches 96; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGCTGTGGCGCACACCTGTAATCCCGACTACCTCGGAGGCTGAGGC 1138  
DB 20871 AAAAATCAGCTGGCGGTGGCGGCGACCGCTGTAATCCCGACTACCTCGGAGGCTGAGGC 20812  
QY 1139 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGGTGACAGCCTCATCA 1193  
DB 20811 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGAGCTAGATACCGCCA 20757

RESULT 12  
US-09-146-053-3/C  
Sequence 3, Application US/09146053A  
Patent No. 639349  
GENERAL INFORMATION:  
APPLICANT: Ryan, James W.  
APPLICANT: Sprinkle, Terry Joe Curtis  
APPLICANT: Venema, Richard C  
TITLE OF INVENTION: Human Amino-peptidase P Gene  
FILE REFERENCE: MCG103  
CURRENT APPLICATION NUMBER: US/09/146,053A  
CURRENT FILING DATE: 1998-09-02  
EARLIER APPLICATION NUMBER: 60/057,854  
EARLIER FILING DATE: 1997-09-02  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3  
LENGTH: 50000  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-146-053-3

Query Match  
Best Local Similarity 90.9%; Pred. No. 7.2e-11;  
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1079 AGAAGACGACGCGCTGTGGCGCACACCTGTAATCCCGACTACCTCGGAGGCTGAGGC 1138  
DB 8689 AAAAATTAGCCCAAGGCTGTGGCGACCGCTGTAATCCCGACTACCTCGGAGGCTGAGGC 8630  
QY 1139 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGGTG 1177  
DB 8629 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGCAGTG 8591

RESULT 13  
US-08-068-945A-1  
Sequence 1, Application US/08068945A  
Patent No. 5616483  
GENERAL INFORMATION:  
APPLICANT: Bjursell, Gunnar  
APPLICANT: Carlsson, Peter  
APPLICANT: Enerback, Sven  
APPLICANT: Hansson, Lemnart  
APPLICANT: Lidberg, Ulf  
APPLICANT: Nilsson, Jeanette  
APPLICANT: Tornell, Jan  
TITLE OF INVENTION: New DNA Sequences  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/068,945A  
FILING DATE: 27-MAY-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9201809-2  
FILING DATE: 11-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9201826-6  
FILING DATE: 12-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9202088-2  
FILING DATE: 03-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9300902-5  
FILING DATE: 19-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Steiner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-052  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)819-8783  
TELEFAX: (212)354-8113  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11531 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
TISSUE TYPE: Mammary gland  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707  
LOCATION: 4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: 8521, 8719..8922, 10124..10321, 10650..11394)  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707  
LOCATION: 4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: 8521, 8719..8922, 10124..10321, 10650..11391)  
OTHER INFORMATION: /EC\_number= 3.1.1.1  
FEATURE:  
NAME/KEY: OTHER INFORMATION: /product= "Bile Salt-Scimulated Lipase"  
LOCATION: 1..1640  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1611..1617  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1641..1727  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4071..4221  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4307..4429  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4707..4904  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6193..6323  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6501..6608  
FEATURE:  
NAME/KEY: exon  
LOCATION: 8335..8420

LOCATION: 6751..6868  
FEATURE:  
NAME/KEY: exon  
LOCATION: 8335..8521  
FEATURE:  
NAME/KEY: exon  
LOCATION: 8719..8922  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10124..10321  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10650..11490  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 11491..11531  
US-08-068-945A-1  
Query Match 3.5%; Score 84.2; DB 1; Length 11531;  
Best local Similarity 91.8%; Pred. No. 4.2e-11;  
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1081 AAAGCAGCCAGCGGTGGTGGCCGACACTGTATCCAGCTACTCGGAGGCTGAGGCAG 1140  
DB 5361 AATTAGCCAGCGGTGGTGGCCCTTGCTGTATCCAGCTACTCGGAGGCTGAGGCAG 5420  
QY 1141 GAGATCGCTTGAACCCGGAGCGGAGGTTGTGCTG 1177  
DB 5421 GAGATCGCTTGAACCTCAGGAGCGGAGGTTGCGGTG 5457  
RESULT 14  
US-08-442-806-1  
Sequence 1, Application US/08442806  
Patent No. 5716817  
GENERAL INFORMATION:  
APPLICANT: Bjursell, Gunnar  
APPLICANT: Carlsson, Peter  
APPLICANT: Enerback, Sven  
APPLICANT: Hansson, Lennart  
APPLICANT: Lidberg, Ulf  
APPLICANT: Nilsson, Jeanette  
APPLICANT: Tornell, Jan  
TITLE OF INVENTION: Genomic DNA Sequences  
TITLE OF INVENTION: Encoding Human BSSL/CEL  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSER: White & Case  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: United States  
ZIP: 10036-2787  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/442,806  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/068,945  
FILING DATE: 27-MAY-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9201809-2  
FILING DATE: 11-JUN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9201826-6  
FILING DATE: 12-JUN-1992  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9202088-2  
FILING DATE: 03-JUL-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9300902-5  
FILING DATE: 19-MAR-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Steiner, Richard J.  
REGISTRATION NUMBER: 35,372  
REFERENCE/DOCKET NUMBER: 1103326-052  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 819-8783  
TELEFAX: (212) 354-8113  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11531 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
TISSUE TYPE: Mammary gland  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707  
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707  
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)  
OTHER INFORMATION: /BC\_number= 3.1.1.1  
OTHER INFORMATION: /product= "Bile Salt-Stimulated Lipase"  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..1640  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1611..1617  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1641..1727  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4071..4221  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4307..4429  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4707..4904  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6193..6323  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6501..6608  
FEATURE:  
NAME/KEY: exon  
LOCATION: 6751..6868  
FEATURE:  
NAME/KEY: exon  
LOCATION: 8335..8521  
FEATURE:  
NAME/KEY: exon  
LOCATION: 8719..8922  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10124..10321  
FEATURE:  
NAME/KEY: exon  
LOCATION: 10650..11490

FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 11491..11531  
US-08-442-806-1  
Query Match 3.5%; Score 84.2; DB 1; Length 11531;  
Best Local Similarity 91.8%; Pred. No. 4.2e-11;  
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1081 AAAGACCCAGGCGGTGTGGCCACACCTGTATCCAGCTACTGTGGAGGCTGAGGCG 1140  
DB 5361 AAATGACGACGCGGTGGCGCTTCTGTATCCAGCTACTGTGGAGGCTGAGGCG 5420  
QY 1141 GAGATCGCTTGACCCGAGGCGGAGGCTTGTGCTG 1177  
DB 5421 GAGATCGCTTGACTCAGAGGCGGAGGCTTGTGCTG 5457  
RESULT 15  
US-09-355-295B-1  
Sequence 1, Application US/09355295B  
Patent No. 6525241  
GENERAL INFORMATION:  
APPLICANT: Dairymple, M.  
APPLICANT: Lundberg, L.  
APPLICANT: Stromqvist, M.  
TITLE OF INVENTION: Expression Methods  
FILE REFERENCE: 1754 SEQUENCE LISTING v2a.txt  
CURRENT APPLICATION NUMBER: US/09/355,295B  
PRIOR FILING DATE: 1999-07-18  
PRIOR APPLICATION NUMBER: PCT/SE99/00648  
PRIOR FILING DATE: 1999-04-21  
NUMBER OF SEQ ID NOS: 5  
SEQ ID NO: 1  
LENGTH: 11531  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genomic DNA  
FEATURE:  
OTHER INFORMATION: Mammary gland source  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707  
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)  
NAME/KEY: mat\_peptide  
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707  
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335  
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)  
OTHER INFORMATION: /BC\_number= 3.1.1.1  
OTHER INFORMATION: /product= "Bile Salt-Stimulated Lipase"  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..1640  
FEATURE:  
NAME/KEY: TATA\_signal  
LOCATION: 1611..1617  
FEATURE:  
NAME/KEY: exon  
LOCATION: 1641..1727  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4071..4221  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4307..4429  
FEATURE:  
NAME/KEY: exon  
LOCATION: 4707..4904  
FEATURE:  
NAME/KEY: exon

```

/ LOCATION: 6193..6323
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 6501..6608
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 6751..6868
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 8335..8521
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 8719..8922
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 10124..10321
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 10650..11490
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: 11491..11531
US-09-355-295B-1

```

```

Query Match 3.5%; Score 84.2; DB 4; Length 11531;
Best Local Similarity 91.8%; Pred. No. 4.2e-11;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

```

QY 1081 AAGCAGCCAGGCGTGCGCAGACCTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 1140
    |||||
DB 5361 AATTAGCCAGCGGTGCGCTGTGCTGTAAATCCAGCTACTCGGAGGCTGAGGCAG 5420
    |||||

QY 1141 GAGATCGCTTGAAACCCGAGGCGAGGCTTGCGTG 1177
    |||||
DB 5421 GAGATCGCTTGAACTCAGGAGGCGAGGCTTGCGTG 5457
    |||||

```

Search completed: March 26, 2004, 05:11:11  
 Job time : 173.702 sec

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 04:26:28 ; Search time 809.174 Seconds

(without alignments)  
11175.829 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429

Sequence: 1 caggagagcccccacagcctca.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table:

IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications NA.\*  
2: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1244.2	51.2	1263	10	US-09-746-783-3
2	587.8	24.2	1717	15	US-10-264-237-259
3	350.4	14.4	1326	13	US-10-071-766-10
4	338.2	13.9	2500	15	US-10-108-260A-1299
5	213	8.8	620	10	US-09-814-351-19504
6	189.2	7.8	1292	14	US-10-198-846-11070
7	175	7.2	1198	9	US-09-880-107-2174
8	161.4	6.6	578	15	US-10-027-632-134530
9	149.4	6.2	748	15	US-10-027-632-134529
10	135.2	5.6	539	14	US-10-029-386-7583
11	134.2	5.5	235	14	US-10-029-386-21283
12	126.8	5.2	1350	9	US-09-768-826-16
13	116	4.8	853	14	US-10-198-846-11346
14	104	4.3	462	14	US-10-198-846-2561
15	103.6	4.3	3186778	15	US-10-027-632-174961

16	102	4.2	615	14	US-10-198-846-7843	Sequence 7843, Ap
17	98	4.0	507	10	US-09-764-872-195	Sequence 195, App
18	90	3.7	624	15	US-10-027-632-243509	Sequence 243509,
19	90	3.7	684973	9	US-09-263-959-1	Sequence 1, Appl
20	89	3.7	31994	9	US-09-764-860-599	Sequence 599, App
21	89	3.7	31994	9	US-09-764-904-71	Sequence 71, Appl
22	89	3.7	31994	14	US-10-091-548-71	Sequence 599, App
23	89	3.7	31994	14	US-10-074-095-599	Sequence 599, App
24	89	3.7	31994	15	US-10-212-872-559	Sequence 11074, A
25	88.4	3.6	723	15	US-10-027-632-11024	Sequence 11024, A
26	88.4	3.6	727	15	US-10-027-632-148248	Sequence 148248,
27	88.4	3.6	727	15	US-10-027-632-148249	Sequence 148249,
28	88.2	3.6	549	15	US-10-027-632-280638	Sequence 280638,
29	88	3.6	30568	9	US-09-764-877-3851	Sequence 3851, Ap
30	88	3.6	30568	15	US-10-242-515-3851	Sequence 3851, Ap
31	87.8	3.6	281	14	US-10-029-058B-15	Sequence 15, Appl
32	87.8	3.6	625	15	US-10-027-632-55059	Sequence 55059, A
33	87.8	3.6	625	15	US-10-027-632-55060	Sequence 55060, A
34	87.8	3.6	625	15	US-10-027-632-293995	Sequence 293995,
35	87.8	3.6	625	15	US-10-027-632-293996	Sequence 293996,
36	87.8	3.6	895	15	US-10-027-632-882	Sequence 822, App
37	87.8	3.6	895	15	US-10-027-632-883	Sequence 823, App
38	87.8	3.6	38918	14	US-10-017-161-2049	Sequence 2049, Ap
39	87.8	3.6	38918	15	US-10-292-798-1955	Sequence 1695, Ap
40	87.8	3.6	133893	13	US-10-161-510-1	Sequence 1, Appl
41	87.4	3.6	468	15	US-10-027-632-195571	Sequence 195571,
42	87.4	3.6	522	15	US-10-027-632-21091	Sequence 21091, A
43	87.4	3.6	616	15	US-10-027-632-99425	Sequence 99425, A
44	87.4	3.6	4963	9	US-09-764-877-2903	Sequence 2903, Ap
45	87.4	3.6	4963	15	US-10-242-515-2903	Sequence 2903, Ap

#### ALIGNMENTS

RESULT 1  
US-09-746-783-3  
; Sequence 3, Application US/09746783  
; Publication No. US20030044935A1  
; GENERAL INFORMATION:  
; APPLICANT: Jacoby, Kenneth  
; McCoy, John M.  
; LaValle, Edward R.  
; Racie, Lisa A.  
; Treacy, Maurice  
; Spaulding, Vikki  
; Agostino, Michael J.  
; Howes, Steven H.  
; Fechtel, Kim  
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 231  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc.  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: U.S.A.  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/746,783  
; FILING DATE: 21-Dec-2000  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Milasincic, Debra J.  
; REGISTRATION NUMBER: 46,931  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1263 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-746-783-3

Query Match 51.2%; Score 1244.2; DB 10; Length 1263;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 1246; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1181 GAGAGGCTCATCAACCCCTTTGGAGAGATGATGATTTTGAACAACCTGATTTGTC 1240
Db 1 GAGAGGCTCATCAACCCCTTTGGAGAGATGATGATTTTGAACAACCTGATTTGTC 1240
QY 1241 GACAGGAATTTGAGAGTGTCCCTGTTGGCTGTGTGATGATGACACAGAGACCTGCTGG 1300
Db 61 GACAGGAATTTGAGAGTGTCCCTGTTGGCTGTGTGATGATGACACAGAGACCTGCTGG 120
QY 1301 ATGAGCGCGGACATGTAATCTGATTAAGCCGAGCCACAGCCCTTGAACAAGCTGCTTC 1360
Db 121 ATGAGCGCGGACATGTAATCTGATTAAGCCGAGCCACAGCCCTTGAACAAGCTGCTTC 180
QY 1361 GCCCAAGTCCGTGAGGCTCTTTATGGCTCCACCTTCAAGATGAGCTTGAACAAAGAG 1420
Db 181 GCCCAAGTCCGTGAGGCTCTTTATGGCTCCACCTTCAAGATGAGCTTGAACAAAGAG 240
QY 1421 GAGATGAGATTTCCAGCCCATGAGAGAGAGAGAGATGCTCAAGCTGATGATTTGGC 1480
Db 241 GAGATGAGATTTCCAGCCCATGAGAGAGAGAGATGCTCAAGCTGATGATTTGGC 300
QY 1481 CGCTTCTAGGCTCAGTCCCATGATGATCAACCTCCAGAGGCAAACTCAAGAGCAAA 1540
Db 301 CGCTTCTAGGCTCAGTCCCATGATGATCAACCTCCAGAGGCAAACTCAAGAGCAAA 360
QY 1541 CTACGTGTGCCCCAAGAGGAATCCCTTCTCAGAGAGGCTGCCCAAAAACCAAGAGCA 1600
Db 361 CTACGTGTGCCCCAAGAGGAATCCCTTCTCAGAGAGGCTGCCCAAAAACCAAGAGCA 420
QY 1601 GCCAAACAGAAAGTTAGGGGCGCAGAGAGAGAGAGAGAGGCTTGAAGCTTGAAGCTTGAAG 1660
Db 421 GCCAAACAGAAAGTTAGGGGCGCAGAGAGAGAGAGAGAGGCTTGAAGCTTGAAGCTTGAAG 480
QY 1661 GCCTTCAAGTCTGCCCCCATGTATCAGAGGCGCAGGCTAATCAAGTCCCAAGAGCGCC 1720
Db 481 GCCTTCAAGTCTGCCCCCATGTATCAGAGGCGCAGGCTAATCAAGTCCCAAGAGCGCC 540
QY 1721 CTGAGCCCACTCCATGTTCTTCCCTGTAAGCAATCAGGCGCTGAAAGCTTCAAGT 1780
Db 541 CTGAGCCCACTCCATGTTCTTCCCTGTAAGCAATCAGGCGCTGAAAGCTTCAAGT 600
QY 1781 GTCAAGGCTTGAACACCAAGCAAAAGCTTAAAGCTGTGAGTCTGGGGCCAAAGAA 1840
Db 601 GTCAAGGCTTGAACACCAAGCAAAAGCTTAAAGCTGTGAGTCTGGGGCCAAAGAA 660
QY 1841 AGTTTGAATTTGCTCAGAGAGAGAGAGAGAGGCTTATGAGAGCCCAAGATCTCA 1900
Db 661 AGTTTGAATTTGCTCAGAGAGAGAGAGAGGCTTATGAGAGCCCAAGATCTCA 720
QY 1901 GTGAGAGAGAAAATCTGTGAGATTTAACTGACGATATATCCGGAATATCTCA 1960
Db 721 GTGAGAGAGAAAATCTGTGAGATTTAACTGACGATATATCCGGAATATCTCA 780
QY 1961 CTCAGAGAACTTTGGAACAATCAACCAACATCACTCACTCACTCACTCACTCACTCACTCA 2020
Db 781 CTCAGAGAACTTTGGAACAATCAACCAACATCACTCACTCACTCACTCACTCACTCACTCA 840
QY 2021 GATCTTATTTGGGCTTTGGAAAAAGGATGAGGATCTTCTAACTGCTTCTTATG 2080
  
```

```

Db 841 GATCCTTATTTGGGCTTTGAAAAACAGGATGAGACATTTCTTAACCTGCTTCTTAATG 900
QY 2081 GAATGCTTGGCCAGCCAGGTCTCTCACTGTGTGTACACAGAGAGACATGATTCAGTCA 2140
Db 901 GGATGCTTGGCCAGCCAGGTCTCTCACTGTGTGTACACAGAGAGACATGATTCAGTCA 960
QY 2141 CAGCCATACAGCTGTCTCACTGAAAGAGAGTGTCTTCAACAAGCCGTAATCAATGTGTA 2200
Db 961 CAGCCATACAGCTGTCTCACTGAAAGAGAGTGTCTTCAACAAGCCGTAATCAATGTGTA 1020
QY 2201 GCTTAAATGATTAATAATCCAGACTCTTCAAGCTTAAATGCTTAAATTAATAAACT 2260
Db 1021 GCTTAAATGATTAATAATCCAGACTCTTCAAGCTTAAATGCTTAAATTAATAAACT 1080
QY 2261 GTGAAAGCTTGAATGATCAACATTTGAAACATTTAACTCAAGACTCTGATTCAGAGTGGGA 2320
Db 1081 GTGAAAGCTTGAATGATCAACATTTGAAACATTTAACTCAAGACTCTGATTCAGAGTGGGA 1140
QY 2321 ACCCTTATGTTCTATCTGAATCCAGACAGCCACACTTAATGATTAATGCTTAAATGA 2380
Db 1141 ACCCTTATGTTCTATCTGAATCCAGACAGCCACACTTAATGATTAATGCTTAAATGA 1200
QY 2381 GTTAAATTAATTCAAATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 2429
Db 1201 GTTAAATTAATTCAAATCTGTTAAAAAATTTAAAAAATTTAAAAAATTTAAAAA 1249

RESULT 2
US-10-264-237-259
/ Sequence 259, Application US/10264237
/ Publication No. US2004009491A1
/ GENERAL INFORMATION:
/ APPLICANT: Biorse et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PAL31P1
/ CURRENT APPLICATION NUMBER: US/10/264,237
/ PRIOR APPLICATION NUMBER: 2002-10-04
/ PRIOR FILING DATE: 2001-05-18
/ PRIOR APPLICATION NUMBER: PCT/US01/16450
/ PRIOR FILING DATE: 2000-05-19
/ NUMBER OF SEQ ID NOS: 2876
/ SOFTWARE: PatentIn Ver. 3.1
/ SEQ ID NO 259
/ LENGTH: 1717
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1072)..(1072)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1596)..(1596)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1634)..(1634)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1657)..(1657)
/ OTHER INFORMATION: n equals a,t,g, or c
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1702)..(1702)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc.feature
/ LOCATION: (1712)..(1712)
/ OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-259
  
```



Query Match 24.2%; Score 587.8; DB 15; Length 1717;  
Best Local Similarity 98.8%; Pred. No. 2.3e-151;  
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1445 GAGGACGAGAGAGATCTCAACGCTGGCATCATTTGGCCGCTTCTTAGGCTGCAATGCCAT 1504  
DB 1 GAGGACGAGAGAGATCTCAACGCTGGCATCATTTGGCCGCTTCTTAGGCTGCAATGCCAT 60  
QY 1505 GATTCACATCTCTCCAGGGGCAACCTCAAGGACCAACTAATGAGGCGCCCAAGAGGGAATCC 1564  
DB 61 GATTCACATCTCTCCAGGGGCAACCTCAAGGACCAACTAATGAGGCGCCCAAGAGGGAATCC 120  
QY 1565 CTCTTCACAGAGGCGCTGCGCAAAAACAAGGCGAGCAACAGAACTTTAGGGGCGAG 1624  
DB 121 CTCTTCACAGAGGCGCTGCGCAAAAACAAGGCGAGCAACAGAACTTTAGGGGCGAG 179  
QY 1625 GAAAGCAACAGAGGCGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGTAT 1684  
DB 180 GAAAGCAACAGAGGCGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGTAT 239  
QY 1685 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTTC 1744  
DB 240 CAGAGGCGAGGCTACTACAGTGGCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTTC 299  
QY 1745 CCCCTAGAACCATCAAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCAATAGACCAACAAAGAC 1804  
DB 300 CCCCTAGAACCATCAAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCAATAGACCAACAAAGAC 359  
QY 1805 AAAAGCTTAAAGCTGTGAGTCTGCGGCGCAAGAAAAGTTTGAATTTGCTTCAGAGAC 1864  
DB 360 AAAAGCTTAAAGCTGTGAGTCTGCGGCGCAAGAAAAGTTTGAATTTGCTTCAGAGAC 419  
QY 1865 GATGGGGCTTATGAGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAAGTGTGAGTTT 1924  
DB 420 GATGGGGCTTATGAGAGACCCAGAAAGTATCTCAAGTGAAGAGAAAAGTGTGAGTTT 479  
QY 1925 AACCTGACGGATATGCGAGAGATCCCGGAAAATCACTTCAAGAACTTTGGAAACATCA 1984  
DB 480 AACCTGACGGATATGCGAGAGATCCCGGAAAATCACTTCAAGAACTTTGGAAACATCA 539  
QY 1985 CCAACCAACATACACACTACACTCAAAAGTCAATGATCTTATTTGAGCCCTTGAAGAAC 2044  
DB 540 CCAACCAACATACACACTACACTCAAAAGTCAATGATCTTATTTGAGCCCTTGAAGAAC 599  
QY 2045 AGGGATG 2051  
DB 600 AGGTCGTG 606

RESULT 3  
US-10-071-766-10/c  
; Sequence 10, Application US/10071766  
; Publication No. US20020192678A1  
; GENERAL INFORMATION:  
; APPLICANT: Huel-Mei Chen  
; TITLE OF INVENTION: GENES EXPRESSED IN SENSESCENCE  
; FILE REFERENCE: PA-0043 US  
; CURRENT APPLICATION NUMBER: US/10/071,766  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PERL Program  
; SEQ ID NO 10  
; LENGTH: 1326  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2  
US-10-071-766-10

Query Match 14.4%; Score 350.4; DB 13; Length 1336;  
Best Local Similarity 99.7%; Pred. No. 7e-86;  
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2053 AGCAGATCTCTAACGCTGCTTCTTAATGGGATGCTTGGCAGCAGAGTCTCAGCTGTG 2112  
DB 1326 AGCAGATCTCTAACGCTGCTTCTTAATGGGATGCTTGGCAGCAGAGTCTCAGCTGTG 1267  
QY 2113 GATACACAGCAGAGACATGATCAAGTCAAGCAGATACAGCTGTCCAGCTGAAGAACTG 2172  
DB 1266 GATACACAGCAGAGACATGATCAAGTCAAGCAGATACAGCTGTCCAGCTGAAGAACTG 1207  
QY 2173 TCTTACACAGCCTGAATGAATGGTGTAGCTTAAATGATTAATAATCCAGACTACTTCAG 2232  
DB 1206 TCTTACACAGCCTGAATGAATGGTGTAGCTTAAATGATTAATAATCCAGACTACTTCAG 1147  
QY 2233 CCTTAATGCTTTTATTCATTAATAAATGTAAGAGTAGAGCAATGGAACATTT 2292  
DB 1146 CCTTAATGCTTTTATTCATTAATAAATGTAAGAGTAGAGCAATGGAACATTT 1087  
QY 2293 AACTCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGAATCCAGACAGCA 2352  
DB 1086 AACTCAGACTGTGATTCAGAGTGGGAACCTTATGTTCTATCTGAATCCAGACAGCA 1027  
QY 2353 CACTTAGATATCTGCCCAACTAATGATTTAATTAATCAAAATCTGTT 2404  
DB 1026 CACTTAGATATCTGCCCAACTAATGATTTAATTAATCAAAATCTGTT 975

RESULT 4  
US-10-108-260A-1299  
; Sequence 1299, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: HI-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1299  
; LENGTH: 2500  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-1299

Query Match 13.9%; Score 338.2; DB 15; Length 2500;  
Best Local Similarity 61.8%; Pred. No. 2.3e-82;  
Matches 538; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 105 ATGACCATCATCTTACACAAAGGAGGCTAATGCGCTTAGGCTCCTTCCCGCGCTG 164  
DB 229 ATGACCTGTACTTACTCTCCAGTAAAGTAGCAATGCACTTTTGGATTTTCAATAGGTTA 288  
QY 165 CTGCTGTGCTGGCGGGGAGCATCTACAAAGCTGTATATGCGAGTTCTTAACTTCTCTG 224  
DB 289 CTCTCAAGTGAAGAGGAGCATCTACAAACTACTGTACAGGGAATTTATGTTTGTGCT 348  
QY 225 CTCTGCTACTACATATCCGCTTATTTATATGAGCTGGCCCTCAAGGAGAAACACAGCTG 284  
DB 349 GTTCTTATACAGCAATAAAGTTGGTGTACAGATTTTACTTACAGAGATCCAAAAACGT 408  
QY 285 ATGTTTGAAGAACTGACTGTGATTTGCGAGCTATCATCAGCTCATGCCATTTCTCTTC 344  
DB 409 TACTTGAAGAAATTTATCAATTTTACTGTGACAGATATGCTGAACAAATTCAGTAACCTTT 468  
QY 345 GTGCTGGCTTCTAGAGTACGCTGTGTGTGACCCGCTGTGTGAACCAATAGAGAACTGTG 404  
DB 469 GTGCTGGCTTCTAGAGTACGCTGTGTGTGACCCGCTGTGTGAACCAATAGAGAACTGTG 528  
QY 405 CCGTGCCCGAGCCGCTCATGAGCTGTGTGTGAGCTTGTGTGAAGGCAAGACGAGCA 464  
DB 529 CCGTGCCCGAGAGGCTATATGTTCTCTCATCTCTTACAGATGTTTCAAGGAGCAAGGAGC 588  
QY 465 GCGCGGCTGTGCGGCGAGCGCTCATCCGCTACGCAACCTGGGCAAGCTGTCTATCTGTG 524

Db 589 GGGGGCTCGTTAGAAAGACGCTGATGCGCTACGTCAATCTTCACTTCCCTGCTCACTTT 648  
 QY 525 CGCAGCGCTGCACACCGCAGTCTTACAGCGCTTCCCCGCGCCAGCACTGTGTGCAAGCA 594  
 Db 649 CGCTCGGTGAGCACTGCTGTGTATCAAAAAGATTTCCCAATGAGACCACTGTGTGAAGCA 708  
 QY 585 GGCCTTATGACTCCGCGAGAACACAGACGTTGGAGAACTGAGCCTACACACAACATG 644  
 Db 709 GGGTTTATGACACAGATGAAAGGAATTTATTCAACCACTCAAGTCTCTCATCTGAAA 768  
 QY 645 TTCTGGGAGCCCTGGAGTGTGGTTTGGCCACCTGTCAATGAAAGCGGTGGCTTGAAGTCCA 704  
 Db 769 TATTGGGTTCAATTCATCTGTGTTTGGAAATCTTGCAACTAAAGCCCGGAATGAAAGGTAGA 828  
 QY 705 ATCCGGGACCCATATCTCTCCAGAGCTGCTGAACGAGATGAACACTTGGCCACTCAG 764  
 Db 829 ATCAGACAGAGTGATATCTGCATTCATTTGATGACTGAAAGAAATCGATACCGCTTTGG 888  
 QY 765 TGTGACACCTGTATGCGCTACGACTGATATGATATCCCACTGATGTATACACAGGTGTC 824  
 Db 889 TGCACCTCTTATTTGGCTTATGACGTGGGTGGGANTCCGCTGGTTTACACCAAGGTGTC 948  
 QY 825 ACTGTGCGGTGTACAGCTTCTTCTGTGACTGTGTCTATGTTGGCGCGCACTTTCTGAACCA 884  
 Db 949 AATCTTGCTGTATATACCTTCTTCTTCCGTGCTGATTTGAACCGCAATTTTTGATCCC 1008  
 QY 885 GCCAAGGCTTACCTTGGCCATGAGCGTGAAGCCGTGTGAGCCGATTTGACGTTCTGTGCA 944  
 Db 1009 ACCAAGGCTAGCGAGGCGCATACCTTGAATCTTTACATTTCCATTTCAACCTCCTACA 1068  
 QY 945 TTCTTCTTATGTTGGCTGCTGAGGTGG 975  
 Db 1069 TTCTTCTTATGACAGGATGGCTTAAAGTAG 1099

```

RESULT 5
US-09-814-353-19504/c
/ Sequence 19504, Application US/09814353
/ Publication No. US20030165831a1
/ GENERAL INFORMATION:
/ APPLICANT: Lee, John
/ APPLICANT: Thompson, Pamela
/ APPLICANT: Lillie, James
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
/ FILE REFERENCE: MRI-006B
/ CURRENT APPLICATION NUMBER: US/09/814,353
/ CURRENT FILING DATE: 2001-03-21
/ PRIOR APPLICATION NUMBER: US 60/191,031
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: US 60/207,124
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: US 60/211,940
/ PRIOR FILING DATE: 2000-06-15
/ PRIOR APPLICATION NUMBER: US 60/216,820
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: US 60/220,661
/ PRIOR FILING DATE: 2000-07-25
/ PRIOR APPLICATION NUMBER: US 60/257,672
/ PRIOR FILING DATE: 2000-12-21
/ NUMBER OF SEQ ID NOS: 22037
/ SOFTWARE: PatSeq for Windows Version 4.0
/ SEQ ID NO 19504
/ LENGTH: 620
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1
/ OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

```

Query Match	8.8%;	Score 213;	DB 10;	Length 620;
Best Local Similarity	66.4%;	Pred. No. 46-48;		
Matches 306;	Conservative	0;	Mismatches 155;	Indels 0;
			Gaps	0;
QY	103	CCATGACACATCACTTACACAGGCCAAGTGGCTTAATGCCCGCTTAAGCTCCTTCGCCGCC	162	
Db	464	CCATGACGGTTTTCATACACTCTCAAGTGGCGAGGCCCGCTTCGAGGTTCTCTGGCC	405	
QY	163	TGCTGCTGTGCTGGCGGGCGAGCATTTACAAGCTGTATATGACGAGTTCTTAATCTTCC	222	
Db	404	TGCTTTCGCCGTGAGGGAGAGATTTACAAGTCTCTTACAAGGAAATTCCTCTCTTTG	345	
QY	223	TGCTGTGCTACTACATCATTCGCTTATTTATAGGCTGGCCCTCACGGAAAGAACACAGC	282	
Db	344	GGGCTTGTAACGGCTGTGCTTAGCATCACTACCGAGCTGTGCTGTGACCCAGAGACAGGT	285	
QY	283	TGAATTTGAGAAACTGACTCTGATTTGGCAGCAGTACATCCAGCTCATCTCCATTCTCT	342	
Db	284	ACGTGTAAGCTCAGAGTGCCCGGTAATCGAACCGCTACAGACACTCATTCCTCTGTCT	225	
QY	343	TCGTGCTGGGCTTCTACGTACCGTGTGCTGTACCCGCTGGTGGAAACAGTACGAAACC	402	
Db	224	TTGTAATGGGTTTCTATGTATGACTCTCTGTGTGAAACGCTGTGTGTCTCCAGTATACAAAGA	165	
QY	403	TGCCGTGGCCCCGACCGGCTCATGAGCTGTGTGTGGGCTTTCGTGAAGGACAGACGAGC	462	
Db	164	TCGGCTGCCAGACCACTGATATGTGGCTCATCTGGGCTACGCTGACCCGGGTGAGACGAGC	105	
QY	463	AAGCGCGCTGCTGTGGGGCGACGCTCATTCGGTACCGCAACTGGGCGAAAGTCTCTATCC	522	
Db	104	GGGGCGGCTCTGCTGGCGGACCTCTCATTCGGCTACGGGAACCTGGCTTGTGTGGTGC	45	
QY	523	TGGCGAGCGTCAGCACCGAGCTTACAAAGCGCTTCCCAAGC	563	
Db	44	TGCGCTGGTCAAGCACCGCGGTGCTTAAGCGCTTCCCAAC	4	

```

US-10-198-846-11070 RESULT 6
; Sequence 11070, Application US/10198846
; Publication NO. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Iallie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/0/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1287, 1288, 1289, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

```

	Query Match	7.8%	Score 189.2	DB 14	Length 1292
Best Local Similarity	62.4%	Pred. No. 2.2e-41			
Matches 296	Conservative	0	Mismatches 178	Indels 0	Gaps 0
QY	349	TTGGCCTTCTACGTAGCGCTGATGTCGATCCGCTGTGTGAACAGTACGAGAACCTTGCCT			

Db 318 TAGGGTTTATGTTACTGTGTAGTGAACCGATGTGGAAACGATTGTGAATTTGCCCT 317  
Qy 409 GGGCCGACCGCCCTCATGAGCTGTGTGCGGGCTTTCGCGAAGGCGAA3GACGACGAGCC 468  
Db 378 GGGCAGACAGGCTTAATGTTCTCTATCTTAAGAGTGTTCACGAGCGACGACGAGGCT 437  
Qy 469 GGGCTGTGGGCGCAGCTCATCGGTACGCCAACCTGTGGCGAACGTTCTATCTCTCGCA 528  
Db 438 GCGTGTTAAGAAGACGCTGATGCGCTGATGCAATCTCACCTCCGCTCATCTTTCCGCT 497  
Qy 529 GCGTCAGCACCGGAGCTGTACAGCGCTTCCCGCCCGCCGACGACCTGTGTGACGAGGCT 588  
Db 498 CCGTGAGCACTGTGTGTACAAAGATTTCCCAATGTGACCGTGTGTGAAGCAGGT 557  
Qy 589 TTATGACTCCGGAGAACCAAGCAGTTGGAGAACTGAGCCATGACACCAATATTTCT 648  
Db 558 TTATGACAAACAGTGAAGAAATTTATCAACCACTCAAGTCTCTCATCTGAATATTT 617  
Qy 649 GGGTCCCTGGGTGTGTGCGCAACCTGTCAATGAGGCGTGTGGAGGTGCAATCC 708  
Db 618 GGGTTCATTCACTCTGTGTGTGAAATCTTGCAACTAAAGCCGGATGAAAGTGAATCA 677  
Qy 709 GGGACCTTATCTCTCTCCAGAGCTGTCTGAACGAGATGAACACCTTCCGTACTCAGTGT 768  
Db 678 GAGACAGTGTGTGATCTGCAATCATGTATGATGATGAAATGAAATGATACCGCTTGTGTGA 737  
Qy 769 GACACCTGTATGCTACGATGATGATGATGATGATGATGATGATGATGATGATGATG 822  
Db 738 GCGCTTATATTCGATATGATGATGATGATGATGATGATGATGATGATGATGATG 791

## RESULT 7

US-09-880-107-2174/c  
Sequence 2174, Application US/09880107  
Patent No. US20020142981A1  
GENERAL INFORMATION:  
APPLICANT: Horne, Darci T.  
APPLICANT: Vockley, Joseph G.  
APPLICANT: Scherf, Uwe  
APPLICANT: Gene Logic, Inc.  
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
FILE REFERENCE: 44921-5028-WO  
CURRENT APPLICATION NUMBER: US/09/880,107  
CURRENT FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/211,379  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: US 60/237,054  
PRIOR FILING DATE: 2000-10-02  
NUMBER OF SEQ ID NOS: 3950  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2174  
LENGTH: 1198  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L:0941  
US-09-880-107-2174

Query Match 7.2%; Score 175; DB 9; Length 1198;  
Best Local Similarity 100.0%; Pred. No. 1.7e-37;  
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2230 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAAGCTAGACTGAACATTGGAAACA 2289  
Db 1198 CAGCCTTAAATGCTTTTATTCATATAAACTGTGAAAGCTAGACTGAACATTGGAAACA 1139  
Qy 2290 TTTAATCAGAGCTGTGATTCAGAGTCGGGAAACCTTAAGTTCTATCTGATCCAAAGACAG 2349  
Db 1138 TTTAATCAGAGCTGTGATTCAGAGTCGGGAAACCTTAAGTTCTATCTGATCCAAAGACAG 1079  
Qy 2350 CCACACCTTAGTACTGCGCCAAACCTAATGATGATTTAATAAATCAAAATCTCGTT 2404  
Db 1078 CCACACCTTAGTACTGCGCCAAACCTAATGATGATTTAATAAATCAAAATCTCGTT 1024

## RESULT 8

US-10-027-632-134530/c  
Sequence 134530, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 134530  
LENGTH: 578  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-134530

Query Match 6.6%; Score 161.4; DB 15; Length 578;  
Best Local Similarity 99.4%; Pred. No. 6.7e-34;  
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 582 GCAGGCTTATGACTCTCGGCAAGACAGACAGAGCTGTGAGAAATGAGCTTACACACAC 641  
Db 183 GCAGGCTTATGACTCTCGGCAAGACAGACAGAGCTGTGAGAAATGAGCTTACACACAC 124  
Qy 642 ATGTTCTGGGTGCTCGGCTGTGTGTTGGCAACCTGTCAATGAAGCGTGTGAGGT 701  
Db 123 ATGTTCTGGGTGCTCGGCTGTGTGTTGGCAACCTGTCAATGAAGCGTGTGAGGT 64  
Qy 702 CGAATCCGGGACCTATCTGCTCCAGAGCTGTGTAACGAGA 744  
Db 63 CGAATCCGGGACCTATCTGCTCCAGAGCTGTGTAACGAGA 21

## RESULT 9

US-10-027-632-134529/c  
Sequence 134529, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358  
 PRIOR FILING DATE: 1999-09-28  
 PRIOR APPLICATION NUMBER: US 60/146,002  
 PRIOR FILING DATE: 1999-08-09  
 NUMBER OF SEQ ID NOS: 325720  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO: 134529  
 LENGTH: 748  
 TYPE: DNA  
 ORGANISM: Human  
 US-10-027-632-134529

Query Match 6.2%; Score 149.4; DB 15; Length 748;  
 Best Local Similarity 90.9%; Pred. No. 1.6e-30;  
 Matches 170; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 558 CCCAGCGCCGACCTGCTGTCGACGAGCCTTTATGACTCCGCGACAAACACAGCAGTTG 617  
 DB 712 CCCAGCGCTTTTTCACCTCCACTCTGACAGCTTTATGACTCCGCGACAAACAGCAG-TG 654  
 QY 618 GAGAACTGAGCCTTACCAACAAATGTTCTGGGTGCTGCTGGGTGCTTGGCAACCTG 677  
 DB 653 GAGAACTGAGCCTTACCAACAAATGTTCTGGGTGCTGCTGGGTGCTTGGCAACCTG 694  
 QY 678 TCAATGAGGCGTGTGCTGAGGTCGCAATCCGAGACCTTATCTGCTCCAGAGCCTGCTG 737  
 DB 593 TCAATGAGGCGTGTGCTGAGGTCGCAATCCGAGACCTTATCTGCTCCAGAGCCTGCTG 534  
 QY 738 AACGAGA 744  
 DB 533 AACGAGA 527

RESULT 10  
 US-10-029-386-7583/c  
 Sequence 7583; Application US/10029386  
 Publication No. US20030194704A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Hanzel, David K.  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
 FILE REFERENCE: A60MICA-X-2  
 CURRENT APPLICATION NUMBER: US/10/029,386  
 CURRENT FILING DATE: 2001-12-20  
 NUMBER OF SEQ ID NOS: 34286  
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 7583  
 LENGTH: 539  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: MAP TO CHR19.3  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
 OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALU 9.00e-56  
 OTHER INFORMATION: NT HIT: g15304094, EVALU 1.00e-130  
 OTHER INFORMATION: SWISSPROT HIT: O76090, EVALU 2.00e-25  
 US-10-029-386-7583

Query Match 5.6%; Score 135.2; DB 14; Length 539;  
 Best Local Similarity 73.3%; Pred. No. 1.1e-26;  
 Matches 173; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 331 GCGTTTACGTCGAGCTGTGTCGACCGCTGTGTGTAACCACTGAGAACTGCGCTGG 410  
 DB 362 GCGTTTATGTGACCGCTGTGTGTCGACCGCTGTGTGTAACCACTGAGAACTGCGCTGG 303  
 QY 411 CCCAGCGCCTCATGAGCCTGTGTGTGCTTCTCTCAAGCAAGAGCAAGAGCGCG 470

DB 302 CCCAGCGCCTCATGTCGTCGTGTGTGCGCGACCGGTGACGAGCCGACGACCGCGCCG 243  
 QY 471 CTGCTGCGCGCAGCCTCATGTCGTCGTGTGTGCGCGACCGGTGACGAGCCGACGAGCGCG 530  
 DB 242 CTCTACCGCGCAGCCTCATGTCGTCGTGTGTGCGCGACCGGTGACGAGCCGACGAGCGCTCC 183  
 QY 531 GTGAGCAACCGCAAGTCGTCGAGCGCTTCCGACCGCGACCGCACTGTGCAAGCAG 586  
 DB 182 GTGAGCAACCGCGAGTGTGTCGAGCGCTTCCGACCGCACTGTGCAAGCAGCAGGCTGG 127

RESULT 11  
 US-10-029-386-21283/c  
 Sequence 21283; Application US/10029386  
 Publication No. US20030194704A1  
 GENERAL INFORMATION:  
 APPLICANT: Penn, Sharon G.  
 APPLICANT: Hanzel, David K.  
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
 FILE REFERENCE: A60MICA-X-2  
 CURRENT APPLICATION NUMBER: US/10/029,386  
 CURRENT FILING DATE: 2001-12-20  
 NUMBER OF SEQ ID NOS: 34288  
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
 SEQ ID NO: 21283  
 LENGTH: 235  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE: MAP TO CHR19.3  
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
 OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALU 4.00e-56  
 OTHER INFORMATION: NT HIT: g18923136, EVALU 1.00e-129  
 OTHER INFORMATION: SWISSPROT HIT: O76090, EVALU 1.00e-24  
 US-10-029-386-21283

Query Match 5.5%; Score 134.2; DB 14; Length 235;  
 Best Local Similarity 73.2%; Pred. No. 1.3e-26;  
 Matches 172; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 352 GCGTTTACGTCGAGCCTGTGTCGTCGACCGCGCTGTGTAACCACTGAGAACTGCGCTGGC 411  
 DB 235 GCGTTTATGTGACCGCTGTGTGTGTAACCGCTGTGTGAGCACTGATGATGCGCGCTGC 176  
 QY 412 CGAGCGCCTCATGAGCCTGTGTGTGCGGCTTGTGTAAGCAAGAGCAAGAGCGCGGC 471  
 DB 175 CGAGCGCCTCATGAGCCTGTGTGTGCGGCAACCGGTGCAAGAGCAAGAGCGCGCGCC 116  
 QY 472 TCGTGGCGGCAAGCCTCATGCGCTTCCGAGCGCCAGCACTGTGTGCAAGCAG 531  
 DB 115 TCAACGGGCGACACTCATGCGCTTACGAGGAGCTCTCGGCGTGTCTATCTGCGCTCG 56  
 QY 532 TCGAGCAACCGAGCTCAAGCGCTTCCGAGCGCCAGCACTGTGTGCAAGCAG 586  
 DB 55 TCGAGCAACCGAGCTTCAAGCGCTTCCGAGCATGAGCACTGTGTGAGGCTGG 1

RESULT 12  
 US-09-768-826-16  
 Sequence 16; Application US/09768826  
 Patent No. US20020012966A1  
 GENERAL INFORMATION:  
 APPLICANT: Shi et al.  
 TITLE OF INVENTION: 18 human secreted proteins  
 FILE REFERENCE: PFS12P1  
 CURRENT APPLICATION NUMBER: US/09/768,826  
 CURRENT FILING DATE: 2001-01-25  
 PRIOR APPLICATION NUMBER: PCT/US00/22350

```

? PRIOR FILING DATE: 2000-08-15
? PRIOR APPLICATION NUMBER: 60/148,759
? PRIOR FILING DATE: 1999-08-16
? NUMBER OF SEQ ID NOS: 61
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 16
? LENGTH: 1350
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (1135)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1148)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1166)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1174)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1181)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1209)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1229)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1266)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1285)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1287)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1290)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1295)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1305)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1324)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1339)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1341)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1343)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1345)
? OTHER INFORMATION: n equals a,t,g, or c
? NAME/KEY: SITE
? LOCATION: (1348)
? OTHER INFORMATION: n equals a,t,g, or c
? US-09-768-826-16
Query Match 5.2%; Score 126.8; DB 9; Length 1350;
Best Local Similarity 69.9%; Pred. No. 3,8e-24;
Matches 186; Conservative 0; Mismatches 77; Indels 3; Gaps 1

```

QY	1174	GGTGGCAGAGAGGCTCATCAACCCCTTTGGAGAGAGATGATGATATTTGAGACCACTG	1233
Db	436	GGTGGCTGACAGATCATCAACCCATTTGGTGAAGATGATGACACTTTGAGCAAAATCA	495
QY	1234	GATTGTGCACAGAAATTTCAGAGTGTCCCTGTTGGCTGTGGATGATGCACACGAGACT	1293
Db	496	GCTCATAGACCGCAACTTTCAGAGTGTCCCTGCTATCCGTGGACGAAATGTACGAAAGCT	555
QY	1294	GCCCTGGATGAGGCCCGGACATGTACTGTGAATTAACCCGAGGCACAGCCCCCTTACAC---	1356
Db	556	TCCCCCGCTGAGAAGACACAGTACTGGAGATGAGAGCACGCGACCACTTACACTGT	615
QY	1351	AGCTGCTTCCGCGCCAGATTCGCTCAGAGCTCCTTTATGGGCTCCACTTCAACATCAGCCT	1410
Db	616	GGCCACGGCGGCGGAGTCTCTGGGGCCCTCATTTCTTGGGCTCCCACTTCAACCTGGCAT	675
QY	1411	GAACCAAGAGGAGATGAGTTCCAGC	1436
Db	676	GAGCGACGACCCTGAGCAGAGCCTTGC	701

```

RESULT 13
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Liilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11346
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
; LOCATION: 43, 50, 51, 54, 57, 853
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346

```

Query Match	4.8%	Score 116;	DB 14;	Length 853;
Best Local Similarity	68.2%	Pred. No. 2.8e-21;		
Matches 161; Conservative	0;	Mismatches 75;	Indels 0;	Gaps 0

Qy	350	GGGCTCTTAAGTGAAGCGCTGGTGTGTAACCGCGGTGGGAAACAGTACAGAAACCTGGCCGTG	409
Db	272	GGGGTTTTTATGTACTCTGGTAGTGAACCGATGGTGGAAACAGTTTGTGAATTTGGCCCTG	3312
Qy	410	GCCCGACCCGCTCATGAGCGCTGGTGCAGGAGTTCGTGCAAGGCAAGGACGACAGGCGC	4659
Db	332	GCCAGACAGGCTAAATGTTCTCATCTCTTAGCAGATGTTTACGGAAAGCGACAGACAGGCGC	3919
Qy	470	GCTGCTGCGGCGCAGCGCTCATCGCTACGCCAACCTGGGCAACGTGCTCATCTGCGCAG	5259
Db	392	CTGCTTAAAGAGACGGCTGATGCGCTCATGTAATCTCACTCCCTGCTCATCTTTGCTC	4515
Qy	530	CGTCAGACACCGAGTCTACAAAGCGTTCTCCCAAGCGCCACAGACCTGTGTCAAGAG	585
Db	452	GGTAGACACTGTGTGTGTAACAAAGATTTCCACATAGGACCACTGTTGTGAAGG	507

## RESULT 14

```

US-10-198-846-2561/c
; Sequence 2561, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2561
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 390..462
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561

```

```

Query Match      4.3%; Score 104; DB 14; Length 462;
Best Local Similarity 67.6%; Pred. No. 4.2e-18;
Matches 146; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

QY 349 TGGGCTTCTACGACGCTGCTGCTGACCCGCTGCTGGAACCACTACAGAACCTGCGCT 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 255 TAGGTTTATGTTACTCTGTTAGTGAACGATGTGGAACAGTTTGTGATTTGGCTT 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 409 GGCCCGACCGCTCATGAGCTGTGTGGGCTTCTGCAAGGCAAGACGAGCAAGCC 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 195 GGCCGACAGGCTATGTTCTCTCATCTCTACAGTGTTCACGAGGCAAGACGAGCC 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 469 GGCTGCTCGGCGGCGCTCATCCGCTACGCAACTGGGCAAGCTGCTCATCTGCGCA 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 135 GCTCTTGAAGAGAGCGCTGATGGCTACGTCATCTACCTCCCTGCTCATCTTTCGCT 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 529 GCGTCAGACCGCAGTCTTACAGCGCTTCCCGACGG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 75 CGGTGAGACCTGCTGTGTAACCTCGGCGCCGACCGCG 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-10-027-632-174961/c
; Sequence 174961, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

```

```

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174961
; LENGTH: 3186778
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3186778)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961

```

```

Query Match      4.3%; Score 103.6; DB 15; Length 3186778;
Best Local Similarity 77.5%; Pred. No. 5.8e-16;
Matches 138; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

```

```

QY 418 GCTCATGAGCTGTGTGTGCGGCTTGTGGAAGCAAGGACGAGCAAGCGGCTGCTC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285842 GCTCTGAGCCTGTGTGTC-OGCTCGCCGAGGCAAGGAGGAGGACGCGGCTGAAG 285784
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 GGCGCAGCTCATCGCTACGCCAACCTGCGCAACGCTCATCTGCGCAGGCTCAGCA 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285783 GAGCATGCTCTCTGCTTGGCCAAACGCGGCAACGTCATTAACTGCGGAAACGCA 285724
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 CCGCATCTACAAAGCGTTTCCCGAGCGCCCAAGCACTGTGTGCAAGAGGCTTATGAC 595
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 285723 CCTCGAACAAAGCGTTTCCACAGCGCGGAGGACACAGGTGCAACAGATGTGAC 285666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Search completed: March 26, 2004, 11:54:40
Job time : 815.507 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 22:13:32 ; Search time 134.65 Seconds  
(without alignments)  
7896.639 Million cell updates/sec

Title: US-09-622-964-28

Perfect score: 1916  
Sequence: 1 gfgccagcagcagcactatca.....aaaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/prodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.2	2.7	1361	4 US-09-489-847-64	Sequence 64, Appl
2	51.8	2.7	1931	4 US-09-019-942-2	Sequence 2, Appl
3	51.8	2.7	1931	4 US-09-099-041A-1	Sequence 1, Appl
4	51.8	2.7	1931	4 US-09-245-281-1	Sequence 1, Appl
5	51.8	2.7	1931	4 US-09-470-271-2	Sequence 2, Appl
6	51.8	2.7	1931	4 US-09-207-359B-1	Sequence 1, Appl
7	51.8	2.7	1931	4 US-09-340-620A-1	Sequence 1, Appl
8	51.8	2.7	1931	4 US-09-865-364-1	Sequence 1, Appl
9	51.8	2.7	1931	4 US-09-748-537-2	Sequence 1, Appl
10	51.4	2.7	2619	4 US-08-983-502-17	Sequence 17, Appl
11	51.4	2.7	2619	4 US-09-516-747-17	Sequence 17, Appl
12	51.4	2.7	2619	5 PCT-US86-10521-17	Sequence 17, Appl
13	51.4	2.7	2887	4 US-08-983-502-14	Sequence 14, Appl
14	51.4	2.7	2887	4 US-09-516-747-14	Sequence 14, Appl
15	51.4	2.7	2887	5 PCT-US96-10521-14	Sequence 14, Appl
16	51	2.7	2060	4 US-09-345-473B-5	Sequence 5, Appl
17	50.8	2.7	1446	4 US-09-593-359-1	Sequence 1, Appl
18	50.8	2.6	1582	3 US-08-545-196B-10	Sequence 10, Appl
19	50.6	2.6	1582	3 US-08-545-196B-12	Sequence 12, Appl
20	48.6	2.5	1937	4 US-09-647-143-1	Sequence 11, Appl
21	48.4	2.5	1584	4 US-09-807-258-11	Sequence 11, Appl
22	47.8	2.5	2053	4 US-09-023-655-854	Sequence 854, App
23	47.6	2.5	1046	1 US-08-361-467B-4	Sequence 4, Appl
24	47.6	2.5	1046	1 US-08-484-332C-4	Sequence 4, Appl
25	47.4	2.5	458	1 US-08-524-757-1	Sequence 24, Appl
26	47.4	2.5	1493	1 US-08-340-820-24	Sequence 24, Appl
27	47.4	2.5	1493	1 US-08-593-535-24	Sequence 24, Appl

28	47.2	2.5	780	2 US-08-540-804-36	Sequence 36, Appl
29	47.2	2.5	780	3 US-08-590-399-36	Sequence 36, Appl
30	47.2	2.5	1952	1 US-08-333-358-1	Sequence 1, Appl
31	47.2	2.5	1952	1 US-08-463-694-1	Sequence 1, Appl
32	47.2	2.5	1952	1 US-08-694-501-1	Sequence 1, Appl
33	47.2	2.5	3214	1 US-08-484-105-17	Sequence 17, Appl
34	47.2	2.5	3214	1 US-08-484-106-17	Sequence 17, Appl
35	47.2	2.5	3722	4 US-10-164-595-9	Sequence 9, Appl
36	47.2	2.5	3862	4 US-10-164-595-5	Sequence 5, Appl
37	47.2	2.5	3937	4 US-10-164-595-7	Sequence 7, Appl
38	47.2	2.5	3985	4 US-10-164-595-3	Sequence 3, Appl
39	47.2	2.5	4895	3 US-09-053-866-1	Sequence 1, Appl
40	47.2	2.5	4895	4 US-09-479-130-1	Sequence 1, Appl
41	47.2	2.5	4895	4 US-09-472-130A-1	Sequence 1, Appl
42	47	2.5	467	2 US-08-841-349-18	Sequence 18, Appl
43	47	2.5	467	4 US-09-431-184A-18	Sequence 18, Appl
44	47	2.5	1313	4 US-09-149-476-112	Sequence 112, App
45	47	2.5	1835	4 US-09-564-808-5	Sequence 5, Appl

## ALIGNMENTS

```
RESULT 1
US-09-489-847-64
; Sequence 64, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 1361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-64

Query Match          2.7%; Score 52.2; DB 4; Length 1361;
Best Local Similarity 71.1%; Pred. No. 0.00069;
Matches 69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY      1820 CTGATTTTACATCTTCCCAACTAGAGCTTTAATTAACGTGAATTTCTTTAGGT 1879
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB      1216 CAGTTGTGAAATATATATTAGCAACTCATGAAATTAACCCCATTAATATTATTATTA 1275
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY      1880 GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB      1276 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1312
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 2
US-09-019-942-2
; Sequence 2, Application US/09019942
; Patent No. 6033855
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
```

```

Query Match: 2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATATTTTTCAGATCTTTCCCAAACTAAGAGCTTTAATAAAGTGAATATTCCT 1873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1825 AAAAGCATGTAAAGTGACTGCTGTTTTCAGAAAGAAATGTGTTTCATAAAGATATTTATAA 1884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1874 TTAGTGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-245-281-1
; Sequence 1, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)...(1833)
US-09-245-281-1

Query Match: 2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATATTTTTCAGATCTTTCCCAAACTAAGAGCTTTAATAAAGTGAATATTCCT 1873
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1825 AAAAGCATGTAAAGTGACTGCTGTTTTCAGAAAGAAATGTGTTTCATAAAGATATTTATAA 1884
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1874 TTAGTGTAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-470-271-2
; Sequence 2, Application US/09470271
; Patent No. 6410689
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
; TITLE OF INVENTION: DOMAIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
;

```



```
/
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows 95
/ SOFTWARE: FastSeq for Windows Version 2.0b
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/470,271
/ FILING DATE:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/019,942
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Melkielejohn, Ph.D., Anita L.
/ REGISTRATION NUMBER: 35,283
/ REFERENCE/DOCKET NUMBER: 07334/068001
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/542-5070
/ TELEFAX: 617/542-8906
/ TELEX: 200154
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1931 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
US-09-470-271-2
```

```
Query Match      2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY      1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAATAACGTAATATCTT 1873
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1825 AAAGCATGTAACTGACTGTTTTCAGAGAAAGAAATGTGTTCTAATAAGATATTATTA 1884

QY      1874 TTAGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1916
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 6
US-09-207-359B-1
/ Sequence 1, Application US/09207359B
/ Patent No. 6469140
/ GENERAL INFORMATION:
/ APPLICANT: Bertin, John
/ TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
/ FILE REFERENCE: 07334-112001
/ CURRENT APPLICATION NUMBER: US/09/207,359B
/ CURRENT FILING DATE: 1998-12-08
/ PRIOR APPLICATION NUMBER: US 09/099,041
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: US 09/019,942
/ PRIOR FILING DATE: 1998-02-06
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (214)...(1833)
US-09-207-359B-1
```

```
Query Match      2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY      1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAATAACGTAATATCTT 1873
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1825 AAAGCATGTAACTGACTGTTTTCAGAGAAAGAAATGTGTTCTAATAAGATATTATTA 1884
```

```
QY      1874 TTAGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1916
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 7
US-09-340-620A-1
/ Sequence 1, Application US/09340620A
/ Patent No. 6482933
/ GENERAL INFORMATION:
/ APPLICANT: Bertin, John
/ TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
/ FILE REFERENCE: 07334-124001
/ CURRENT APPLICATION NUMBER: US/09/340,620A
/ CURRENT FILING DATE: 1999-06-28
/ PRIOR APPLICATION NUMBER: US 09/245,281
/ PRIOR FILING DATE: 1999-02-05
/ PRIOR APPLICATION NUMBER: US 09/207,359
/ PRIOR FILING DATE: 1998-12-08
/ PRIOR APPLICATION NUMBER: US 09/099,041
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: US 09/019,942
/ PRIOR FILING DATE: 1998-02-06
/ NUMBER OF SEQ ID NOS: 71
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (214)...(1833)
US-09-340-620A-1
```

```
Query Match      2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY      1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAATAACGTAATATCTT 1873
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1825 AAAGCATGTAACTGACTGTTTTCAGAGAAAGAAATGTGTTCTAATAAGATATTATTA 1884

QY      1874 TTAGTGAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1916
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 8
US-09-865-364-1
/ Sequence 1, Application US/09865364
/ Patent No. 6613521
/ GENERAL INFORMATION:
/ APPLICANT: Bertin, John
/ TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
/ FILE REFERENCE: 07334-112001
/ CURRENT APPLICATION NUMBER: US/09/865,364
/ CURRENT FILING DATE: 2001-05-25
/ PRIOR APPLICATION NUMBER: US 09/207,359
/ PRIOR FILING DATE: 1998-12-08
/ PRIOR APPLICATION NUMBER: US 09/099,041
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: US 09/019,942
/ PRIOR FILING DATE: 1998-02-06
/ NUMBER OF SEQ ID NOS: 47
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 1931
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (214)...(1833)
```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
;

```

```
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
;
```

```
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/516,747
/ FILING DATE: 01-Mar-2000
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/983,502
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/ APPLICATION NUMBER: IL 117,932
/ FILING DATE: 16-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Browdy, Roger L.
/ REGISTRATION NUMBER: 25,618
/ REFERENCE/DOCKET NUMBER: WALLACH=19
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-5197
/ TELEFAX: (202) 737-3528
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2619 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-516-747-17

Query Match          2.7%; Score 51.4; DB 4; Length 2619;
Best Local Similarity 72.0%; Pred. No. 0.0015;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1824 ATTTAGCATCTTCCCAACTAGAGTTTATTAACGTGAATCTTTTAGTGAAA 1883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2527 ATTTAGTATCTTTTCCAAACATTTTATAGATAAATATATTGATCTTAAA 2586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
PCT-US96-10521-17
/ Sequence 17, Application PC/TUS9610521
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
/ TITLE OF INVENTION: AND OTHER PROTEINS
/ NUMBER OF SEQUENCES: 34
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/10521
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ PRIOR APPLICATION DATA:
```

```
/
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 117,932
/ FILING DATE: 16-APR-1996
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2619 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ PCT-US96-10521-17

Query Match          2.7%; Score 51.4; DB 5; Length 2619;
Best Local Similarity 72.0%; Pred. No. 0.0015;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1824 ATTTAGCATCTTCCCAACTAGAGTTTATTAACGTGAATCTTTTAGTGAAA 1883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2527 ATTTAGTATCTTTTCCAAACATTTTATAGATAAATATATTGATCTTAAA 2586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-983-502-14
/ Sequence 14, Application US/08983502
/ Patent No. 6399327
/ GENERAL INFORMATION:
/ APPLICANT: David WALLACH
/ APPLICANT: Mark P. BOLDIN
/ APPLICANT: Tanya M. GONCHAROV
/ APPLICANT: Yuri V. GOLTSSEV
/ TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
/ TITLE OF INVENTION: AND OTHER PROTEINS
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Browdy and Neimark
/ STREET: 419 Seventh Street N.W., Ste. 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20004
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ OPERATING SYSTEM: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/983,502
/ FILING DATE: 16-JAN-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/10521
/ FILING DATE: 14-JUN-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/ PRIOR APPLICATION DATA:
```

ATTORNEY/AGENT INFORMATION:  
NAME: Browdy, Roger L.

MOLECULE TYPE: CDNA  
PCT-US96-10521-14

Query Match	2.7%	Score 51.4;	DB 5;	Length 2887;
Best Local Similarity	72.0%;	Pred. No. 0.0016;		
Matches 67; Conservative	0;	Mismatches 26;	Indels 0;	Gaps 0;

Qy	1824	ATTTACGATCTTCCCAACTAAGAGTTTAATAACGTAATATTCCTTTAGTGAAA	1883
Db	2791	ATTTAAGTATCTTTTCCAAAACATTTTAAATAGATTAATATTAATTGATCTTAA	2850
Qy	1884	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	1916
Db	2851	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	2883

Search completed: March 26, 2004, 05:11:13  
 Job time : 136.65 secs

**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 04:26:28 ; Search time 638.278 Seconds  
(without alignments)  
11175.829 Million cell updates/sec

Title: US-09-622-964-28

Perfect score: 1916  
Sequence: 1 ggcacagcagcatgactatca.....aaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Prod. No. is the number of results predicted by chance a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	458.2	23.9	2500	US-10-108-260A-1299	Sequence 1299, Ap
2	227.6	11.9	1263	US-09-746-783-3	Sequence 3, Appl1
3	194.4	10.1	620	US-09-814-353-19504	Sequence 19504, A
4	187.8	9.8	1350	US-09-768-826-16	Sequence 16, Appl1
5	181.2	9.5	1282	US-10-198-846-11070	Sequence 11070, A
6	125.2	6.5	578	US-10-027-632-134530	Sequence 134530, A
7	122.2	6.4	539	US-10-029-386-7583	Sequence 7583, Ap
8	119.8	6.3	235	US-10-029-386-21283	Sequence 21283, Ap
9	114.4	6.0	853	US-10-198-846-11346	Sequence 11346, A
10	113.2	5.9	748	US-10-027-632-134529	Sequence 134529, A
11	94.4	4.9	462	US-10-198-846-2561	Sequence 2561, Ap
12	94.4	4.9	507	US-10-198-846-7843	Sequence 7843, Ap
13	94.2	4.9	507	US-09-764-872-195	Sequence 195, App
14	75.8	4.0	3186778	US-10-027-632-174961	Sequence 174961, A
15	69	3.6	1717	US-10-264-237-259	Sequence 259, App

16	65	3.4	65	10	US-09-908-975-29941	Sequence 29941, A
17	63.6	3.3	442	10	US-09-814-353-14343	Sequence 14343, A
18	63	3.3	751	14	US-10-198-846-1862	Sequence 1862, Ap
19	62.2	3.2	260	10	US-09-814-353-1601	Sequence 1601, Ap
20	62.2	3.2	260	10	US-09-814-353-7959	Sequence 7959, Ap
21	61.6	3.2	298	10	US-09-814-353-806	Sequence 806, App
22	61.6	3.2	298	10	US-09-814-353-7178	Sequence 7178, Ap
23	60.4	3.2	463	10	US-09-814-353-13563	Sequence 13563, A
24	56.2	2.9	3490	9	US-09-925-301-44	Sequence 44, Appl1
25	55.6	2.9	440	12	US-10-424-599-5477	Sequence 5477, Ap
26	54.6	2.8	644	9	US-09-984-245-84	Sequence 84, Appl1
27	54.6	2.8	644	10	US-09-966-262-84	Sequence 84, Appl1
28	54.6	2.8	644	10	US-09-983-966-84	Sequence 84, Appl1
29	54.6	2.8	644	14	US-10-143-090-84	Sequence 84, Appl1
30	54.6	2.8	1465	15	US-10-264-237-468	Sequence 468, App
31	53	2.8	128	14	US-10-066-543-1216	Sequence 1216, Ap
32	53	2.8	473	12	US-10-424-599-12926	Sequence 12926, A
33	52.8	2.8	380	10	US-09-814-353-17491	Sequence 17491, A
34	52.4	2.7	3637	9	US-09-729-674-3	Sequence 3, Appl1
35	52.2	2.7	1361	12	US-10-351-334-64	Sequence 64, Appl1
36	52	2.7	338	15	US-10-131-827-8601	Sequence 8601, Ap
37	52	2.7	444	12	US-10-424-599-109078	Sequence 109078, A
38	51.8	2.7	1331	9	US-09-748-537-2	Sequence 2, Appl1
39	51.8	2.7	1331	9	US-09-966-262-84	Sequence 1, Appl1
40	51.8	2.7	1331	13	US-10-133-780-2	Sequence 2, Appl1
41	51.8	2.7	1331	13	US-10-105-931-1	Sequence 1, Appl1
42	51.8	2.7	1331	13	US-10-118-964-1	Sequence 1, Appl1
43	51.8	2.7	1331	14	US-10-295-981-1	Sequence 1, Appl1
44	51.6	2.7	273	12	US-10-424-599-33887	Sequence 33887, A
45	51.6	2.7	840	9	US-09-925-301-498	Sequence 498, App

ALIGNMENTS

RESULT 1  
US-10-108-260A-1299  
; Sequence 1299, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: NO. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1299  
; LENGTH: 2500  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-108-260A-1299

Query Match 23.9%; Score 458.2; DB 15; Length 2500;  
Best local Similarity 63.5%; Pred. No. 2,5e-118;  
Matches 700; Conservative 0; Mismatches 403; Indels 0; Gaps 0;

QY	11	ATGACTATACCTTACCAACAAGTAGCCGATGCCGCTGCTTCGTCCTC	70
DB	229	ATGACTGTCACTTACTCCAGTAAGAGCAAACTTTTGGATTTCATAGATTA	288
QY	71	CTCCCTGTGCTGCGAGCAGCATCTCAAGCTGCTATGAGAAATTCCTGCTCATA	130
DB	289	CTCCCTCACTGAGAGCAGCATCTCAAGCTGCTATGAGAAATTCCTGCTCATA	348
QY	131	TTCTCTTACTATTCATCCGCTGACTCTACAGATGTTCTCTGAGTATCAGACCTG	190
DB	349	GTTCTTTATACAGCATATAGTTTGGTTCACATTTCTTACAGAGATCCAAACGT	408
QY	191	TTGTTTGAAGAGCTGCTGTGATCGACAGCTATTCAGTCAATCCCTATATCTTC	250
DB	409	TACTTTGAAAATTTATCAATTTACTGTGACAGATATGCTGAACAAATTCAGTAACCTTT	468

QY 251 GTTCTGGGTTTCTATGTTACATTGNGTGGAGCCGCTGGTGGAGCCAGTACGAGACTTG 310  
DB 469 GTGCTGGGTTTATGTTAATCTGTGTAAGAAACGATGGTGAACCAAGTTTGAATTGG 528  
QY 311 CCGTGGCCCGACCGCCTCATGATCCAGGTGTCTAGCTTGTGGAGGGCAAGATGAGAA 370  
DB 529 CCTGGCCGAGAGAGGTATGTTCCCATCTCTAGCAAGTGTTCAGGAAGCCAGAGCA 588  
QY 371 GGCCTGTTGCTGCGGAGACGCTCATCGGCTACGCCATCTTGAGGCAAGTCTCATCTG 430  
DB 589 GGGCGCTGCTAGAAAGAGACGCTGATGCGCTTCAATCTCACTCCCTGCTCATCTTT 648  
QY 431 CGCAGATACAGACCTCGGCTTCAAGAGCCTTCCACTCTTCAACCACTGCTGATGAGA 490  
DB 649 CGCTGGTGAAGACCTGCTGTGTAACAAATTTCCCAATGAGACCACTGGTGAAGCA 708  
QY 491 GGTATATGACCATGGGAGACATAGACAGTTGACAGAGTTGGGCTTACACACACACA 550  
DB 709 GGTATATGACCAAGATGAAAGAAATATTCACCACTCAAGTCTCTCATCTGAAA 768  
QY 551 TTCTGGGTCCTGGGTGTGTGTTGCGCACTGTCTAATAGAGCCTATCTTGAAGTGA 610  
DB 769 TATTTGGGTTCCATTCTGTGTTGAAATCTTGCAACTAAAGCCGGAATGAGGTAGA 828  
QY 611 ATCCGGAGAACCGTCTGCTCCAGACCTGATGAAAGAGTGTACTTGGTACTGAG 670  
DB 829 ATCAGAGACAGTGTGATCTGCAATCATGATGACGAAATGATGATACCGCTCTTGG 888  
QY 671 TGTGACAGCTGATGCTACAGACCTGATAGATAGATCCCATTTGCTGACACAGTGTG 730  
DB 889 TGCAGCTCTTATTCGCTTATGATCTGGGTGGATTCGCTGGTTTACACCGAGTTGTC 948  
QY 731 ACAGTGGAGTATACAGCTTTTCTTTCGATGCTTGAATGGAGGCACTTTGAAACCA 790  
DB 949 ACTCTGTGTGTATCTCTCTCTTCTTGGCTGCTGATGAGAGCCAGTTTGTGATCCC 1008  
QY 791 AACAGAGCATCCAGGCGCATGAGATGATCTGCTGCTCTTCAACATCTGCA 850  
DB 1009 ACCAAAGGCTACGAGAGGATGACTTGGATCTTAACTTCCATCTTCAACCTCTTCA 1068  
QY 851 TTCTTATCTACATGGGCTGTGCTGAAGTGGCAAGCAAGCTCATCAACCCCTTGGGAG 910  
DB 1069 TTCTTCTTCTATGCAAGATGGCTTAAGTAGAGACAGCTTATCAACCCCTTTGAGAA 1128  
QY 911 GACGATGATGATTTTGAAGCTAACTGATCATTTGAAGAACTGCAAGTGTCTCTG 970  
DB 1129 GATGATGATGATTTTGAATCTGATGATGATGATGATGATGATGATGATGATGAT 1188  
QY 971 TCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1030  
DB 1189 GCTGTGAGCAAAATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1248  
QY 1031 GCAGCCCTTCAGCCGCTTACACAGCTGCTTTCGCAAGTCTGCGCGCATTTCTTCATG 1090  
DB 1249 TCTGCTGCTGCGCCACATACATGAGCTGCTGATGATGATGATGATGATGATGAT 1308  
QY 1091 GGTTCACCTTCAACATCAGCCT 1113  
DB 1309 GGGTCAACAGTCCAGATGGGGCT 1331

RESULT 2  
US-09-746-783-3

Sequence 3, Application US/09746783

Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth

McCoy, John M.

Lavallite, Edward R.

Racie, Lisa A.

Treacy, Maurice

Spaulding, Vikki

Agostino, Michael J.

Howes, Steven H.  
Fechel, Kim  
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES  
ENCODING THEM  
NUMBER OF SEQUENCES: 231  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1263 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-746-783-3

Query Match 11.9%; Score 227.6; DB 10; Length 1263;  
Best Local Similarity 67.7%; Pred. No. 2.6e-53;  
Matches 383; Conservative 0; Mismatches 144; Indels 39; Gaps 3;

QY 884 GAACAGCTCATCAACCCCTTGGGAGAGACGATGATGATTTTGAAGCTAACTGATCAT 943  
DB 1 GACAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAAGCTAACTGATGATGTC 60  
QY 944 GACAGAACTCTGAGGTGTCCTGTTGTCGGATGGATGATGATGATGATGATGATGATGAT 1003  
DB 61 GACAGAACTCTGAGGTGTCCTGTTGTCGGATGGATGATGATGATGATGATGATGATGAT 120  
QY 1004 ATGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1063  
DB 121 ATGAGACCGGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180  
QY 1064 GCCAGGTCTGCGCGGATCTTTCATGAGCTTCACTTCAACATGACCTTAAAGAAAGAA 1123  
DB 181 GCCAGGTCTGCGCGGATCTTTCATGAGCTTCACTTCAACATGACCTTAAAGAAAGAA 240  
QY 1124 GACTTGAAGCTTTGTCAAAGAGAGGCTGACAGGATTAAGAAAGAGTGGCTATAC 1183  
DB 241 GAGATGAGATTCAGGCGCAATGAGAGG-----ACGAGAGAGATGCTACGCT 288  
QY 1184 AGCAGCATAGGCTGCTTCTTAGACTGCAACCCAAACCTACATCTTCCCTTGAAGAGC 1243  
DB 289 GGCATATGTCGCGCTTCTTAGGCTGCAAGTCCATGATCAACATCTTCCAGGGGCAAGC 348  
QY 1244 TTAAAGCAAACTATTTGTTCTTAAAGAA-----CCCTCTCTGGAAGCCAGTGAAG 1297  
DB 349 TCAAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCTTCAAGAGGCGTCCCA 408  
QY 1298 GATGCCAACCAAGAAACCAAGAAAGAT-----GTCTGAAGATT 1336  
DB 409 AACCAAGGACGACCAAGAAAGCTTAAAGGGCCAGAGAAAGCAACAAAGGCTTGAAGCTT 468



Qy 1337 AAGGCTGTGACTTCTTGAATGTGTCCAGGTTTAAGAGAGAGACTCCATTGTGGC 1396  
Db 469 AAGGCTGTGAGACGCTTCAAGTGTGCCCCCACTGTATCAGAGCGCAAGCTACAGTGGC 528  
Qy 1397 CCACAGGACCCACGACGACCCCTAC 1422  
Db 529 CCACAGACGCGCTCTCAGCCCACTCC 554

## RESULT 3

US-09-814-353-19504/c  
; Sequence 19504, Application US/09814353  
; Publication No. US20030165831A1  
; GENERAL INFORMATION:  
; APPLICANT: Lee, John  
; APPLICANT: Thompson, Pamela  
; APPLICANT: Little, James  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: MRI-0068  
; CURRENT APPLICATION NUMBER: US/09/814,353  
; CURRENT FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: US 60/191,031  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: US 60/207,124  
; PRIOR FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: US 60/211,940  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 60/216,820  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/220,661  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: US 60/257,672  
; PRIOR FILING DATE: 2000-12-21  
; SOFTWARE: FastSeq for windows Version 4.0  
; NUMBER OF SEQ ID NOS: 22037  
; SEQ ID NO 19504  
; LENGTH: 620  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1  
; OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-19504

Query Match 10.1%; Score 194.4; DB 10; Length 620;  
Best Local Similarity 63.9%; Pred. No. 4.1e-44;  
Matches 294; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

Qy 9 CCATGACTATACCTACACAAAGTACGCAATCCCGCTCGGTTGTTCTGCTCC 68  
Db 464 CCATGACGGTTTCATCACTCTCAAGTGGCGGCGCTTCGAGTTCTGCGCC 405  
Qy 69 TCCTCTGCTGCGGAGGAGCATAGAGCTGTGTATGAGAGTTCCTGTTTCA 128  
Db 404 TCGTCTCCGCTGAGAGGAGATACAGCTCTCTACAGAGGATTCCTCTTGG 345  
Qy 129 TATTCCTACTATTCATCCGTGAGCTACAGATGGTCTCTGATGATCAGCAGC 188  
Db 344 GGGGCTGTAGCGTGTAGATCACTACCGGCTGTGCTGACCCAGAGCAGAGT 285  
Qy 189 TGTGTTTGAAGAGCTGCTGTGACTGAGACAGTCAATTGACTTATCTCT 248  
Db 284 AGGTGATCTCAGGTGGCGCGGTACTGCAACCGCTCAGACGCTCATTCCTTGTCT 225  
Qy 249 TCGTTCGGGTTTCTATGTTTACATGGTGTGAGCGCGTGTGAGACGAGTACGAACT 308  
Db 224 TTGTATTTGGGTTTCTATGACTCTGTGTGAGACCGCTGGTGTCCAGTACACAGCA 165  
Qy 309 TCCGCTGCCGACCGCTCATGATCAGGTGTCTAGCTTCTGTGAGAGGCAAGATGAG 368

Db 164 TCCCGCTGCAGACGACGTATGTGCTCATCTCGGCTAGCGGTGACCGCGTGACAC 105  
Qy 369 AAGGCGTTTGTGAGGCGGACGCTCATCCGCTAGGCGCATTCGAGGCGCAAGTGTATCC 428  
Db 104 GGGGCGGCTGTGCGCGGACCTCATCCGCTACGCGAACCCTGGCGTTGTGTGTGC 45  
Qy 429 TGGCAGCATCAGACCTGCTCTACAGAGCGTTTCCAC 468  
Db 44 TGGCTCGGTGACGACCGCGGTCTTACGCGCTTCCCA 5

## RESULT 4

US-09-768-826-16  
; Sequence 16, Application US/09768826  
; Patent No. US20020012966A1  
; GENERAL INFORMATION:  
; APPLICANT: Shi et al.  
; TITLE OF INVENTION: 18 human secreted proteins  
; FILE REFERENCE: PFS12P1  
; CURRENT APPLICATION NUMBER: US/09/768,826  
; CURRENT FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: PCT/US00/22350  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/148,759  
; PRIOR FILING DATE: 1999-08-16  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 16  
; LENGTH: 1350  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1135)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1148)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1166)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1174)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1181)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1209)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1229)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1266)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1285)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1287)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1290)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1295)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1305)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (1324)

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16
```

Query Match 9.8%; Score 187.8; DB 9; Length 1350;  
Best Local Similarity 60.1%; Pred. No. 4,8e-42; Indels 48; Gaps 2;  
Matches 377; Conservative 0; Mismatches 202;

```
544 CAACATCTCTGGGTCCTCTGGTGTGTTGCAACTTGTCAATGAAGCCTATCTTG 603
58 CAACAGTACTGGGTCCTCTGGTGTGTTGCAACTTGTCAATGAAGCCTATCTTG 117
604 AGGTGAATCCGGGACACCGTCTGCTCAAGCTGATGAATGAGTGTGTACTTTGCG 663
118 CGGGGCAATACGTGACGATATCGCTCTGCTCACTTTTGAAGACTCAACAGTACCG 177
664 TACTAGTGTGACAGCTGTATGCCACAGCTGGAATATCCATGTTGTGACACA 723
178 AGCCAGTGTACAGATGATATCCATGATACAGTATGATGATGATGATGATGAT 237
724 GGTGTGACAGTGTGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 783
238 AGTGTGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 297
784 GAACCAACAAAGAGTATCCCAAGGCTGATGATGATGATGATGATGATGATGATGAT 826
298 GAGGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 357
827 -----GTGCTGTCTTCAACATCTGCAATCTTCTTAT 858
358 CCCAGCTCTGGGAGACCCCGGACATGATGATGATGATGATGATGATGATGATGAT 417
859 CTACATGAGCTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
418 CTATGTGTGTGCTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 477
919 TGATTTTGAAGTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
478 CGACTTTGAGACAAATCAGCTATAGACGCACTTGAGTGTCTCTGTATCCGTGGA 537
979 TGGGAGGACCCGAACTTGTCTCCAGTGAAGTGAATGATGATGATGATGATGATGAT 1038
538 CGAAATGTACAGAACTTGTCTCTCCGCTGAGAAAGACACTTGGAGTGAAGACAGCC 597
1039 TCAGCCGCTTACAC--AGTGTCTTCTGCAAGTGTCTGCGGCAATCTTCTTCA 1095
598 GAGGACACCTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 657
1096 CACCTTCAACATCAGCTTAAAGAA 1122
658 CACCTTCAACCTGCGATGAGCGACGA 684
```

RESULT 5  
US-10-198-846-11070  
; Sequence 11070; Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Millie, James

```
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIOR FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1287-1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070
```

Query Match 9.5%; Score 181.2; DB 14; Length 1292;  
Best Local Similarity 61.4%; Pred. No. 3,4e-40;  
Matches 291; Conservative 0; Mismatches 183; Indels 0; Gaps 0;

```
255 TGGGTTTCTATTTTCACTTTGTTGTTGAGCCGCTGTTGAGCCAGTACAGAACTTGGCT 314
318 TGGGTTTCTATTTTCTCTGTTGTTGTTGAGCCGCTGTTGAGCCAGTACAGAACTTGGCT 377
315 GGGCCGACCGCTCAGTATCCAGTGTCTAGCTTGTGTTGAGGAGGAGGAGTGAAGAGCC 374
378 GGGCCGACCGCTCAGTATCCAGTGTCTAGCTTGTGTTGAGGAGGAGGAGTGAAGAGCC 437
375 GTTTGTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 434
438 GCTGTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 497
435 GCATGAGCAGCTCGTGTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 494
498 CGGTGAGCAGCTCGTGTCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 557
495 TTATGACCCATGGGAGCAATAGCAGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 554
558 TTATGACCAACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 617
555 GGGTGGCTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 614
618 GGGTTCATTTCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 677
615 GGGACACCTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 674
678 GAGACAGTGTGATCTGCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 737
675 GACAGCTGATCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
738 GCTCTTATTCGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 791
```

RESULT 6  
US-10-027-632-134530/c  
; Sequence 134530; Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; POLYMORPHISMS IN THE HUMAN GENOME  
; FILE REFERENCE: 108827, 129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 134530  
LENGTH: 578  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-134530

Query Match  
Best Local Similarity 6.5%; Score 125.2; DB 15; Length 578;  
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 488 GCAGGTTTATGACCCATGCGGGAACATAAGCAGTTGCGAGAGTTGG3CTACACACAC 547  
DB 183 GCAGGCTTTATGACTCCGCGAGAACACAGAGTTGAGAACTGAGCCTACACACAC 124  
QY 548 ACATTCTGGGTCCTCGGTCGTCGTTGGCCAACTGTCAATGAGGCTTATCTTGAGGT 607  
DB 123 ATGTCCTGGGTCCTCGGTCGTCGTTGGCCAACTGTCAATGAGGCGTGGAGGT 64  
QY 608 CCAATCCGGGACACCTCTCTCCAGAGCCTGATGAAAGAG 649  
DB 63 CCAATCCGGGACCTTATCTCTCCAGAGCCTGATGAAAGGT 22

RESULT 7  
US-10-029-386-7583/c  
Sequence 7583, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
NUMBER OF SEQ ID NOS: 34286  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 7583  
LENGTH: 539  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR19.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
OTHER INFORMATION: EST\_HUMAN HIT: AA573517.1, EVALUE 9.00e-56  
OTHER INFORMATION: NT HIT: g115304094, EVALUE 1.00e-130  
OTHER INFORMATION: SWISSPROT HIT: 076090, EVALUE 2.00e-25  
US-10-029-386-7583

Query Match  
Best Local Similarity 6.4%; Score 122.2; DB 14; Length 539;  
Matches 169; Conservative 0; Mismatches 78; Indels 0; Gaps 0;  
QY 247 CTTGCTTGGGTTTCTATGATGATGAGTGGAGCCCGCTGAGGAGCAGTACGAGAA 306

DB 372 CTCCTCGAGGCTTTATGACGCTGGTGTGAACCCCTGGTGGAGCCAGCACTGATG 313  
QY 307 CTTGCGGTGGCCCGGACCGCTCATGATCCAGGCTGATGTTGTTGAGAGGCAAGATGA 366  
DB 312 CATGCCGTCGCGGACCGCTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 253  
QY 367 GGAAGCCGTTTGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 426  
DB 252 CCGGCGCCCTCTTACCGGCGACACTCATGCGCTACGAGGCTCTCGGCGCTCAT 193  
QY 427 CTTGCGGACATCAGGACCTCGGTCCTTACAGCGCTTTCCACTTTCACCACTGTCGT 486  
DB 192 CTTGCGGCTCGTCAGGACCGGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 133  
QY 487 AGCAGGT 493  
DB 132 GGTGCT 126

RESULT 8  
US-10-029-386-21283/c  
Sequence 21283, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C  
FILE REFERENCE: AEMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
NUMBER OF SEQ ID NOS: 34286  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 21283  
LENGTH: 235  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO CHR19.3  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2  
OTHER INFORMATION: EST\_HUMAN HIT: AA573517.1, EVALUE 4.00e-56  
OTHER INFORMATION: NT HIT: g18923136, EVALUE 1.00e-129  
OTHER INFORMATION: SWISSPROT HIT: 076090, EVALUE 1.00e-24  
US-10-029-386-21283

Query Match  
Best Local Similarity 6.3%; Score 119.8; DB 14; Length 235;  
Matches 163; Conservative 0; Mismatches 72; Indels 0; Gaps 0;  
QY 258 GTTTCATGTTATGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 317  
DB 235 GCTTTATGTCAGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 176  
QY 318 CCGACCGCTCATGATCCAGGTGTCTAGCTTGTGAGGCGCAAGATGAGAAAGCCGTT 377  
DB 175 CCGAGCGGCTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 116  
QY 378 TCTGCGGCGGACGCTCATCCGTCATGCGTCATGCGTCATGCGTCATGCGTCATGCG 437  
DB 115 TCTACCGGCGGACACTCATGCGTCATGCGTCATGCGTCATGCGTCATGCGTCATGCG 56  
QY 438 TCAGCAGCTCGGTCATACAGCGCTTCCACTTTCACCACTGTCGTCAGCAGG 492  
DB 55 TACGACCGCGGTGTCACAGCGCTTCCCACTATGACCACTGTCGTCAGCAGG 1

RESULT 9  
US-10-198-846-11346  
Sequence 11346, Application US/10198846

```
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinmann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ PRIOR FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 11346
/ LENGTH: 853
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
/ LOCATION: 43, 50, 51, 54, 57, 853
/ OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346
```

```
Query Match      6.0%; Score 114.4; DB 14; Length 853;
Best Local Similarity 67.8%; Pred. No. 2e-21;
Matches 160; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
QY 256 GGGTTCTATGTTACATTGGTGTGAGCCGCTGTGAGAGCACTAGAGACCTTGCCTG 315
DB 272 GGGGTTTATGTTACTGTGTAGTGAACCGATGTGAACTTTGTGATTGGCCCTG 331
QY 316 GCCCGACCCCTCATGATCCAGGTGTCTAGCTTCTGTGAGGCGAAGATGAGAGCCG 375
DB 332 GCCAGACAGGCTATATTTCTCTACTCTAGCAATGTTTCAAGAGGAGAGCAAGCGGG 391
QY 376 TTTGCTGGGCGGACGCTATCCGCTACGCGCATCTCGGGCAAGTGCTCATCTTGGCGAG 435
DB 392 CCGCTTAAGAGGACCGCTGATGCGCTACGCTACATCTCCTCCCTCTCATCTTTGCTTC 451
QY 436 CATGACACCTCGGTCTACAGGCTTCCACTCTTCAACCACTGCTGTAGCAG 491
DB 452 GGTGACACCTGCTGTGTGACAAAAGATTTCACACATGAGCACGATGATGAGGAG 507

RESULT 10
US-10-027-632-134529/C
/ Sequence 134529, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
```

```
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 134529
/ LENGTH: 748
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-134529
```

```
Query Match      5.9%; Score 113.2; DB 15; Length 748;
Best Local Similarity 85.2%; Pred. No. 4.1e-21;
Matches 138; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
```

```
QY 488 GCAGTTTATGACCATGGAGACATAGCAAGTTGCAGAGTTGGGCTTACACACAC 547
DB 688 GCAGGCTTATGATCTCCGCGACAGACAGAG-TCGAAACTAGGCTTACACACAC 630
QY 548 AATTTCTGGTCCCTCGGTGTGTGTTGCTTGCACCTGTCTATGAGGCTTATGAGGT 607
DB 629 ATGTTCTGGTCCCTCGGTGTGTGTTGCTTGCACCTGTCTATGAGGCTTATGAGGT 570
QY 608 CGAATCCGGAGACCGTCTCTCCAGAGCCGTGATGATGAG 649
DB 569 CGAATCCGGAGCCATCTCTGCTCCAGAGCCGTGATGAG 528
```

```
RESULT 11
US-10-198-846-2561/C
/ Sequence 2561, Application US/10198846
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinmann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2561
/ LENGTH: 462
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 390_462
/ OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561
```

```
Query Match      4.9%; Score 94.4; DB 14; Length 462;
Best Local Similarity 67.0%; Pred. No. 6.3e-16;
Matches 134; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 255 TGGGTTTATGTTACATTGTTGTGAGCCGCTGTGAGACGATGACGAACCTTGCCT 314
DB 255 TGGGTTTATGTTACTGTGTAGTGAACCGATGTGGAACCAAGTTTGAATTGGCCCT 196
QY 315 GCCCGACCCGCTCATGATCCAGGTGTCTAGTGTGAGGCGAAGATGAGAGGCG 374
DB 195 GCCAGACAGGCTAATGTTCTCATCTTAGAGATGTTACGGAACGAGACGAGCGGC 136
QY 375 GTTTCGCGGCGGACGCTCATCCGCTACGCCATCTGGGCGCAAGTCTCATCTTGGCGA 434
DB 135 GCTGCTTAAGAGAGAGCGATGCGGTAGTGAATCTACCTCCCTGCTCATCTTGGCT 76
QY 435 GCATGACACCTCGGCTTAC 454
DB 75 CGGTGACACCTGCTGTGTAC 56
```

```

RESULT 12
US-10-198-846-7843
; Sequence 7843, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7843
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-7843

```

Query Match	Similarity	Score	DB	Length
Beet Local	134	67.0%	Conservative	0
Matched	134	67.0%	Conservative	0
Query	255	TGGGTTTCATGCTTACATTGGTGTAGAGCCGCTGTGTGAGCCAGTACGAGAACCTTGCCGT	314	
Db	353	TAGGGGTTTATGTTACTCTGGTAGTGACCGAGTGGTGGAGAACGATTGTGAATTTGGCCT	412	
Qy	315	GGCCCGACCCGCTCATGATCCAGGTGTTAGCTTCGTGGAGGGCAAGATGAGAGAGCC	374	
Db	413	GGCCAGACAGGGCTAATGTTCCCTCATCTTAGCAGTGTTCACGGAGAGAGACACACGGGC	472	
Qy	375	GTTTGCTGGGGGCGACGCTCATCCGCTACGGCCATCCGAGGCAAGTGCATGCTGGCA	434	
Db	473	GCTTGCTTGAAGAGAGCGCTGATGCGCTAGCTCAATCTACCTCCCTGATCTTCGCT	532	
Qy	435	GCATGAGCAGCTCGGTCTAC	454	
Db	533	CGGTGAGCAGCTGCTGTGTAC	552	
RESULT 13				
US-09-764-872-195				
Sequence 195, Application US/09764872				
Publication No. US20030050231A1				
GENERAL INFORMATION:				
APPLICANT: Rosen et al.				
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies				
FILE REFERENCE: PA125				
CURRENT APPLICATION NUMBER: US/09/764,872				
Prior application data removed - consult PALM or file wrapper				
NUMBER OF SEQ ID NOS: 957				
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO 195				
LENGTH: 507				
TYPE: DNA				
ORGANISM: Homo sapiens				
FEATURE:				
NAME/KEY: SITE				
LOCATION: (490)				
OTHER INFORMATION: n equals a,t,g, or c				
US-09-764-872-195				
Query Match	4.9%	Score 94.2	DB 10	Length 507
Beet Local Similarity	61.0%	Pred. No. 7.6e-16		
Matched	134	61.0%	Pred. No. 7.6e-16	

Matches	153;	Conservative	0;	Mismatches	98;	Indels	0;	Gaps	0;
Qy	9	CCATGACTATACCTACTACAAACAAATAGACCAATGCCCGCGCTGGTTCGTTCTCCGCTCC	68						
Db	35	CCATGACGGTTTATATACCTCTCAAAAGGGGGGGCCCGCTTGAGAGTTTCTCTGGCC	94						
Qy	69	TCTCTCTGTGCTGCGAGGAGCACTCTACAAGCTGTGTATAGAGAAATTCCTTGTCTTCA	128						
Db	95	TGCTTCTCCGCTGAGGGGAAGCATCTACAAAGCTCTCTCAACAGAAATTCCTCCCTTTG	154						
Qy	129	TATTCCTCTATATATTCATCCGTGAGACTCTACAGAAATGTTCTCTCGAGTATCAGACGC	188						
Db	155	GGGCGCTTGATGCGCTGTGCTTAGCATCACTTACCGGCTGTCTCTAACCCAGAGAGCAGAGT	214						
Qy	189	TGTTGTTTGAGAAACCTGGCTCTGTACTGCGACAGCTACATTCACTCATCCCTTATCTT	248						
Db	215	ACGTGTATGCTCAGAGTGGCCCGAGTACTGCAACCGCTCAGCAGAACCTCATTTCCCTGTCTCT	274						
Qy	249	TTCGTTCTGGGT	259						
Db	275	TTGTATTGGGT	285						

```

RESULT 14
US-10-027-632-174961/C
; Sequence 174961, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174961
; LENGTH: 3186778
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(3186778)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961

Query Match 4.0%; Score 75.8; DB 15; Length 3186778;
Best Local Similarity 67.6%; Pred. No. 2.4e-08;
Matches 121; Conservative 0; Mismatches 57; Indels 1; Gaps 1;

OY 324 GCCTCATGATCCAGGTGCTCTGAGCTTCGTGAGAGGCGCAAGATAGAGAAAGGCCGTTGCTGC 383
Db 285842 GCCTCTGAGGCTTGATGTC-GGCTTCGCGGAGGCAAGACAGACAGCGCCGCTGMAAG 285784A

OY 384 GAGCGAGCGCTCATCGCTAGCGCAATCTGGGCGCAAGTCTCATCTCTGCGACATCAGCA 443
Db 285783 GAGCAGTCTCTCTCGCTTCGCGCAACGCGCAACGTCTATTAACTTCGGAACGGCAGCA 285724A

OY 444 CCTCGCTTACAGAGCGCTTCCACTTTCACCACTCGATGCTTACAGAGTTTATGACC 502

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 01:53:34 ; Search time 8577 Seconds

(without alignments)  
11198.340 Million call updates/sec

Title: US-09-622-964-1\_COPY\_13900\_16115

Portect score: 2216  
Sequence: 1 attctcctgtcttcttcacg.....ctaatgagtttaataaac 2216

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenEmbl: 1: gb\_ba: \*  
2: gb\_ptg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_ats: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: gb\_da: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_ats: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_hcg\_hum: \*  
31: em\_hcg\_inv: \*  
32: em\_hcg\_other: \*  
33: em\_hcg\_mus: \*  
34: em\_hcg\_pln: \*  
35: em\_hcg\_rtd: \*  
36: em\_hcg\_mam: \*  
37: em\_hcg\_vtc: \*  
38: em\_sy: \*  
39: em\_hcgo\_hum: \*  
40: em\_hcgo\_mus: \*  
41: em\_hcgo\_other: \*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2215	100.0	16125	6	BD136719	BD136719 Best's ma
2	2188.6	98.8	142052	9	AF139813	AF139813 Homo sapi
3	2188.6	98.8	196080	9	AC004228	AC004228 Homo sapi
4	2187	98.7	112309	9	AC003025	AC003025 Homo sapi
5	2187	98.7	163024	9	AP006260	AP006260 Homo sapi
6	2187	98.7	163915	2	AC087451	AC087451 Homo sapi
7	2187	98.7	166867	2	AP003733	AP003733 Homo sapi
8	2185.4	98.6	133683	2	AC084857	AC084857 Homo sapi
9	1608.4	72.6	160169	2	AC0051664	AC0051664 Homo sapi
10	1561.6	70.5	106648	2	AC004588	AC004588 Homo sapi
11	1339.8	60.5	2435	9	BC041664	BC041664 Homo sapi
12	688.4	31.1	706	9	HSWMD2P10	AF073499 Homo sapi
13	662	29.9	726	9	AY357926S10	AY357925 Macaca fa
14	643.8	29.1	1263	6	BD132659	BD132659 Secreted
15	643.8	29.1	2210	9	AF057169	AF057169 Homo sapi
16	643.8	29.1	2420	9	AF057170	AF057170 Homo sapi
17	642.2	29.0	1758	6	AX745964	AX745964 Sequence
18	642.2	29.0	2171	9	AF073501	AF073501 Homo sapi
19	642.2	29.0	2229	6	BD136720	BD136720 Best's ma
20	642.2	29.0	2429	6	BD136721	BD136721 Best's ma
21	642.2	29.0	2441	9	BC015220	BC015220 Homo sapi
22	642.2	27.1	2187	9	AY357925	AY357925 Macaca fa
23	399.4	18.0	424	6	BD059099	BD059099 Secreted
24	389.4	17.6	482	9	HSWMD2P11	AF073500 Homo sapi
25	318.6	14.4	1289	4	AY064707	AY064707 Sub scrof
26	318	14.4	356	6	AX908879	AX908879 Sequence
27	318	14.4	356	6	BD044412	BD044412 Sequence
28	302	13.6	450	9	AY357926S11	AY357926 Macaca fa
29	164	7.4	438	11	G05922	G05922 human STS W
30	164	7.4	1119	9	HUMFERRH	M97164 Human ferrit
31	164	7.4	1198	6	AX409527	AX409527 Sequence
32	164	7.4	1198	6	HUMFERR1TH	L20941 Human ferrit
33	136	6.1	195646	9	AC093709	AC093709 Pan trogl
34	135	6.1	186896	9	AC019057	AC019057 Homo sapi
35	134.4	6.0	181049	9	AC093131	AC093131 Papio anu
36	133.2	6.0	65042	6	AR235846	AR235846 Sequence
37	133.2	6.0	65042	6	AX702088	AX702088 Sequence
38	133.2	6.0	149480	6	AX329775	AX329775 Sequence
39	133.2	6.0	149480	6	AX329776	AX329776 Sequence
40	133.2	6.0	149480	6	AX336193	AX336193 Sequence
41	133.2	6.0	149480	9	HU095740	U95740 Human Chrom
42	133.2	6.0	149480	9	AC130455	AC130455 Homo sapi
43	133.2	6.0	156212	9	HS12684	AL022316 Human DNA
44	133.2	6.0	175839	9	AC004491	AC004491 Homo sapi
45	133.2	6.0	180399	2	AC079252	AC079252 Homo sapi

## ALIGNMENTS

RESULT 1  
BD136719 16125 bp DNA linear PAT 18-SEP-2002  
LOCUS Best's macular dystrophy gene.  
ACCESSION BD136719  
VERSION BD136719.1 GI:23231664  
KEYWORDS JP 2002504559-A/1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 16125)  
AUTHORS Petrushkin, K., Caskey, T.C., Metzker, M. and Wadelius, C.  
TITLE Best's macular dystrophy gene  
JOURNAL Patent: JP 2002504559-A 1 (12-FEB-2002)

Pred. No. is the number of results predicted by chance to have a

COMMENT  
MERCK & CO INC, CLAES WADELIUS  
OS Homo sapiens (human)  
PN JP 2002504559-A/1  
PD 12-FEB-2002  
PF 22-FEB-1999 JP 2000533447  
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI  
KONSTANTIN PETRUKHIN, THOMAS C CASEKEY, MICHAEL METZKER, CLAES PI  
WADELIUS  
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/66// PC  
C12P21/08  
CC C12N5/00, C12N15/00  
CH Best's macular dystrophy gene  
FH Key Location/Qualifiers  
FT source 1..16125  
/organism='Homo sapiens (human)'.  
Location/Qualifiers  
1..16125  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'

ORIGIN  
Query Match 100.0%; Score 2215; DB 6; Length 16125;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATCTCTGTTCTTTCCAGCCTGAAACAAGAGAGATGAGTTCCAGCCCAATCAGAG 60  
DB 13900 ATCTCTGTTCTTTCCAGCCTGAAACAAGAGAGATGAGTTCCAGCCCAATCAGAG 60  
QY 61 ACAGAGAGATGCTCAGCGCTGGCATATTGGCCGCTTCTAGGCTCTGAGTCCCATGATC 120  
DB 13960 ACAGAGAGATGCTCAGCGCTGGCATATTGGCCGCTTCTAGGCTCTGAGTCCCATGATC 120  
QY 121 ACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGGCCCAAGAGGGAATCCCTTC 180  
DB 14020 ACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGGCCCAAGAGGGAATCCCTTC 180  
QY 181 TTCACAGAGGCGCTGCCAAAACAACAAGGCAAGCCAAACAGAACGTTAGGGGCCAGAGAG 240  
DB 14080 TTCACAGAGGCGCTGCCAAAACAACAAGGCAAGCCAAACAGAACGTTAGGGGCCAGAGAG 240  
QY 241 ACAACAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCCCTGTATCA 300  
DB 14140 ACAACAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCCCTGTATCA 300  
QY 301 GGCAGGCTACATCAAGTSCCCCAACAGAGCCCTCAGCCCACTCCCATGTTCTCCCC 360  
DB 14200 GGCAGGCTACATCAAGTSCCCCAACAGAGCCCTCAGCCCACTCCCATGTTCTCCCC 360  
QY 361 TAGAACATCAAGCGCCGTCMAAGCTTCAAGTGTCAAGGCAATGACCAACAAGACAAA 420  
DB 14260 TAGAACATCAAGCGCCGTCMAAGCTTCAAGTGTCAAGGCAATGACCAACAAGACAAA 420  
QY 421 GCTTAAAGACTGTGAGTTCTGGGCGCAAAAAGTTTGAATTGCTCTCAGAGACGATG 480  
DB 14320 GCTTAAAGACTGTGAGTTCTGGGCGCAAAAAGTTTGAATTGCTCTCAGAGACGATG 480  
QY 481 GGGCTTGTATGAGCAACCAAGAGTATCTCAAGTGTGAGGAGAAACTGTGAGTTAAAC 540  
DB 14380 GGGCTTGTATGAGCAACCAAGAGTATCTCAAGTGTGAGGAGAAACTGTGAGTTAAAC 540  
QY 541 TGACGGATGTCGAGAGATCCCGAAGATCAGCTTCAAGAACCTTTGGAACATCAACA 600  
DB 14440 TGACGGATGTCGAGAGATCCCGAAGATCAGCTTCAAGAACCTTTGGAACATCAACA 600  
QY 601 CCAACATACACTACACTCAAGATCAGATGAGATCTTATTGGCCTTGGAAAAACAGT 660  
DB 14500 CCAACATACACTACACTCAAGATCAGATGAGATCTTATTGGCCTTGGAAAAACAGT 660  
QY 661 CTGTCTCCACTGGAACCAAGGCGCACTGCAATTCCTGCTGCCCCCAAGCTTCCCTTG 720  
DB 14560 CTGTCTCCACTGGAACCAAGGCGCACTGCAATTCCTGCTGCCCCCAAGCTTCCCTTG 720

QY 721 CTGTGAGCCTACCTTCTCCACAAATTTCTAGAGTTTCCATCTGCCAGAGCACTGG 780  
DB 14620 CTGTGAGCCTACCTTCTCCACAAATTTCTAGAGTTTCCATCTGCCAGAGCACTGG 780  
QY 781 ACCATCGCCGACACTGGCTTGGGATATATATCTTGGCCACTTTCACAGGATCTTAGGGA 840  
DB 14680 ACCATCGCCGACACTGGCTTGGGATATATATCTTGGCCACTTTCACAGGATCTTAGGGA 840  
QY 841 AGTGTGGGACCTTTTCTACTTCACTCCCTGGATATACCCGGAACCTTTGGGACGAG 900  
DB 14740 AGTGTGGGACCTTTTCTACTTCACTCCCTGGATATACCCGGAACCTTTGGGACGAG 900  
QY 901 GTGAAGAAATAGATGTTGTGCTGACAGAAATGCTGTGAGAACTGCCCAAGGCTGAC 960  
DB 14800 GTGAAGAAATAGATGTTGTGCTGACAGAAATGCTGTGAGAACTGCCCAAGGCTGAC 960  
QY 961 AGGCAAGCTTAGCTGAGAGATGTTATCACTGGGCCCAACTTACTTTAGCAAGGCTGG 1020  
DB 14860 AGGCAAGCTTAGCTGAGAGATGTTATCACTGGGCCCAACTTACTTTAGCAAGGCTGG 1020  
QY 1021 CTGACCCAAACCATAGAGGTGGCACTGACCTGATGACAGATGAACACTTCCCATATAC 1080  
DB 14920 CTGACCCAAACCATAGAGGTGGCACTGACCTGATGACAGATGAACACTTCCCATATAC 1080  
QY 1081 TATTTAGGTTAGTACCCAAAGCACTTACAGGAAAGGCTGGCAGGAACTGCTCTAGG 1140  
DB 14980 TATTTAGGTTAGTACCCAAAGCACTTACAGGAAAGGCTGGCAGGAACTGCTCTAGG 1140  
QY 1141 AACTGTGATGATGAGGTTGAGGGTGTCCAGCGCTTATGATCATTTTCTACCTGCTG 1200  
DB 15040 AACTGTGATGATGAGGTTGAGGGTGTCCAGCGCTTATGATCATTTTCTACCTGCTG 1200  
QY 1201 GGAACCTCACCAAAATCTTCTGCTTCTTGGGCTCAGCCCAAGCTGTCAAAATCA 1260  
DB 15100 GGAACCTCACCAAAATCTTCTGCTTCTTGGGCTCAGCCCAAGCTGTCAAAATCA 1260  
QY 1261 GATATTTCCCTTATTTCCAGATTTCTGACACTGTCAACCAATTAACAACCCACTT 1320  
DB 15160 GATATTTCCCTTATTTCCAGATTTCTGACACTGTCAACCAATTAACAACCCACTT 1320  
QY 1321 CAGCCCAATCAAGTGGAGAGAGTAACTTCTCTTCTGATTTCTCAAGCACTACT 1380  
DB 15220 CAGCCCAATCAAGTGGAGAGAGTAACTTCTCTTCTGATTTCTCAAGCACTACT 1380  
QY 1381 TTCAAGGCTCAAGACGAGACTTATGATGAAACCTTAAAGGGAACAATTTCA 1440  
DB 15280 TTCAAGGCTCAAGACGAGACTTATGATGAAACCTTAAAGGGAACAATTTCA 1440  
QY 1441 CTGCTTTCTAGGCTTACAGAGAACTTGGCAAACTGTGGGCTGTTCAGCAAGATG 1500  
DB 15340 CTGCTTTCTAGGCTTACAGAGAACTTGGCAAACTGTGGGCTGTTCAGCAAGATG 1500  
QY 1501 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGTCA 1560  
DB 15400 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGTCA 1560  
QY 1561 GGAAGTTGAGCAACCTGGGCCAATGATGATAAACCCTATCTTCAACAAAAAATACAA 1620  
DB 15460 GGAAGTTGAGCAACCTGGGCCAATGATGATAAACCCTATCTTCAACAAAAAATACAA 1620  
QY 1621 ATCACTGGCGGTGTGTGTGCTGTAGTCCCAAGCAGAGAGTGAAGGAGAAATTCG 1680  
DB 15520 ATCACTGGCGGTGTGTGTGCTGTAGTCCCAAGCAGAGAGTGAAGGAGAAATTCG 1680  
QY 1681 TTGAACCCAGAGAGTGTGTGTGAGATGAGATTCAGCAATTCAGCTGGGCGAC 1740  
DB 15580 TTGAACCCAGAGAGTGTGTGTGAGATGAGATTCAGCAATTCAGCTGGGCGAC 1740  
QY 1741 GGAAGTGAAGTGTCTCAAAAAAAGAGAGAGTGTCTCAACTTTGGCTCTTACT 1800  
DB 15640 GGAAGTGAAGTGTCTCAAAAAAAGAGAGAGTGTCTCAACTTTGGCTCTTACT 1800



Qy	1801	GCAACATTTTGGATTGTAATGAGTACCTTCATCTACTTAATGCTGTTAATCTTTCAT	1860	/chromosome="11"
				/map="11q13"
		15700	GCAACATTTTGGATTGTAATGAGTACCTTCATCTACTTAATGCTGTTAATCTTTCAT	/clone="pDJ759j12"
				complement(19..1855)
Qy	1861	TTTCACTAGGGAAGACACATTCCTAAGCTGCTTCTTAATGGGAGTCTTCCGACCCA	1920	/rpt_family="L1"
				1861..2001
		15760	TTTCACTAGGGAAGACACATTCCTAAGCTGCTTCTTAATGGGAGTCTTCCGACCCA	/rpt_family="L1"
				1990..2317
Qy	1921	GGTCTCAGCTGTGTGTGACACAGCAGACATGATCAGTCAAGCCATCAAGCTGTCC	1980	/rpt_family="Alu"
				complement(2031..2257)
		15820	GGTCTCAGCTGTGTGTGACACAGCAGACATGATCAGTCAAGCCATCAAGCTGTCC	/rpt_family="SVA"
				2320..2974
				/rpt_family="L1"
				2320..2974
				/rpt_family="Alu"
				2385..3540
				/rpt_family="Alu"
				complement(4034..4124)
Qy	1981	ACAAGTGAAGACGTGTCTCTCAACAGCCTGAATCAATGGTTAGTTATGATTAATAAT	2040	/rpt_family="MIR"
				4417..4995
		15880	ACAAGTGAAGACGTGTCTCTCAACAGCCTGAATCAATGGTTAGTTATGATTAATAAT	/rpt_family="SVA"
				complement(5503..6111)
Qy	2041	CCCAAGTACTTCAAGCTTTAATGCTTTATTCATTAATAAAGCTGTAAGTACTGAA	2100	/rpt_family="SVA"
				complement(5503..6111)
		15940	CCCAAGTACTTCAAGCTTTAATGCTTTATTCATTAATAAAGCTGTAAGTACTGAA	/rpt_family="SVA"
				5569..6095
				/rpt_family="SVA"
				complement(7239..7542)
				/rpt_family="Alu"
				7306..7497
				/rpt_family="SVA"
				7752..7828
				/rpt_family="MIR"
				9099..9403
				/rpt_family="Alu"
				complement(9526..9897)
				/rpt_family="MER1"
				9635..9752
				/rpt_family="MER1B"
				11924..12197
				/rpt_family="Alu"
				complement(11955..12151)
				/rpt_family="SVA"
				complement(12767..13466)
				/rpt_family="Alu"
				16757..17039
				/rpt_family="Alu"
				complement(16775..17010)
				/rpt_family="SVA"
				complement(18220..18507)
				/rpt_family="Alu"
				18225..18485
				/rpt_family="SVA"
				20101..23446
				/gene="HUMFERRITTH"
				/note="HUMAN ferritin heavy chain gene"
				join(20101..20421,22218..22364,22621..22746,22842..23446)
				/gene="HUMFERRITTH"
				/product="ferritin heavy chain"
				complement(21466..21731)
				/rpt_family="Alu"
				21506..21714
				/rpt_family="SVA"
				complement(23272..37375)
				/gene="VMD2"
				/note="Best macular dystrophy gene"
				complement(join(23272..23629,24840..25478,27690..27841,
				28155..28255,29427..29579,30274..30351,30740..30894,
				31787..32020,32537..32631,35777..35964,37308..37375))
				/gene="VMD2"
				/product="Best macular dystrophy protein"
				complement(23735..23954)
				/rpt_family="Alu"
				23789..23956
				/rpt_family="SVA"
FEATURES				
source				
				1..142092
				/organism="Homo sapiens"
				/mol_type="genomic DNA"
				/db_xref="taxon:9606"
REFERENCE				
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17				
18				
19				
20				
21				
22				
23				
24				
25				
26				
27				
28				
29				
30				
31				
32				
33				
34				
35				
36				
37				
38				
39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				
54				
55				
56				
57				
58				
59				
60				
61				
62				
63				
64				
65				
66				
67				
68				
69				
70				
71				
72				
73				
74				
75				
76				
77				
78				
79				
80				
81				
82				
83				
84				
85				
86				
87				
88				
89				
90				
91				
92				
93				
94				
95				
96				
97				
98				
99				
100				
101				
102				
103				
104				
105				
106				
107				
108				
109				
110				
111				
112				
113				
114				
115				
116				
117				
118				
119				
120				
121				
122				
123				
124				
125				
126				
127				
128				
129				
130				
131				
132				
133				
134				
135				
136				
137				
138				
139				
140				
141				
142				
143				
144				
145				
146				
147				
148				
149				
150				
151				
152				
153				
154				
155				
156				
157				
158				
159				
160				
161				
162				
163				
164				
165				
166				
167				
168				
169				
170				
171				
172				
173				
174				
175				
176				
177				
178				
179				
180				
181				
182				
183				
184				
185				
186				
187				
188				
189				
190				
191				
192				
193				
194				
195				
196				
197				
198				
199				
200				
201				
202				
203				
204				
205				
206				
207				
208				
209				
210				
211				
212				
213				
214				
215				
216				
217				
218				
219				
220				
221				
222				
223				
224				
225				
226				
227				
228				
229				
230				
231				
232				
233				
234				
235				
236				
237				
238				
239				
240				
241				
242				
243				
244				
245				
246				
247				
248				
249				
250				
251				
252				
253				
254				
255				
256				
257				
258				
259				
260				
261				
262				
263				
264				
265				
266				
267				
268				
269				
270				
271				
272				
273				
274				
275				
276				
277				
278				
279				
280				
281				
282				
283				
284				
285				
286				
287				
288				
289				
290				
291				
292				
293				
294				
295				
296				
297				
298				
299				
300				
301				
302				
303				
304				
305				
306				
307				
308				
309				
310				
311				
312				
313				
314				
315				
316				
317				
318				
319				
320				
321				
322				
323				

repeat_region	26429. .26707	/rpt_family="Alu"	
repeat_region	complement(26444. .26696)	/rpt_family="SVA"	
repeat_region	27103. .27349	/rpt_family="Alu"	
repeat_region	complement(27106. .27323)	/rpt_family="SVA"	
repeat_region	complement(29161. .29307)	/rpt_family="Alu"	
repeat_region	complement(29762. .30043)	/rpt_family="Alu"	
repeat_region	29786. .30017	/rpt_family="SVA"	
repeat_region	30462. .30516	/rpt_family="MTR"	
repeat_region	complement(31095. .31409)	/rpt_family="Alu"	
repeat_region	33005. .33260	/rpt_family="Alu"	
repeat_region	33247. .33597	/rpt_family="L1MC2"	
repeat_region	complement(33849. .34431)	/rpt_family="Alu"	
repeat_region	34076. .34411	/rpt_family="SVA"	
repeat_region	complement(34640. .34890)	/rpt_family="Alu"	
repeat_region	34972. .35401	/rpt_family="Alu"	
repeat_region	complement(34988. .35392)	/rpt_family="SVA"	
repeat_region	36238. .36487	/rpt_family="Alu"	
repeat_region	complement(36253. .36457)	/rpt_family="SVA"	
repeat_region	complement(38661. .38855)	/rpt_family="L1"	
repeat_region	complement(40997. .41224)	/rpt_family="Alu"	
repeat_region	41056. .41213	/rpt_family="SVA"	
repeat_region	42745. .43020	/rpt_family="Alu"	
repeat_region	complement(42763. .42966)	/rpt_family="SVA"	
repeat_region	43312. .43457	/rpt_family="Alu"	
repeat_region	complement(43963. .44217)	/rpt_family="L1"	
repeat_region	complement(45296. .45533)	/rpt_family="SVA"	
repeat_region	complement(45300. .45401)	/rpt_family="L1R5"	
repeat_region	complement(45407. .45533)	/rpt_family="L1R5"	
repeat_region	complement(45542. .45818)	/rpt_family="Alu"	
repeat_region	45598. .45770	/rpt_family="SVA"	
repeat_region	complement(45862. .46004)	/rpt_family="L1"	
repeat_region	complement(45998. .46282)	/rpt_family="Alu"	
repeat_region	46022. .46260	/rpt_family="SVA"	
repeat_region	complement(46692. .47583)	/rpt_family="Alu"	
repeat_region	46885. .47215	/rpt_family="SVA"	
repeat_region	47363. .47571	/rpt_family="SVA"	
repeat_region	complement(47665. .47773)	/rpt_family="SVA"	

repeat_region	/rpt_family="Alu"	repeat_region	complement(47792. .48077)
repeat_region	/rpt_family="Alu"	repeat_region	complement(48110. .48388)
repeat_region	/rpt_family="L1"	repeat_region	complement(48827. .49102)
repeat_region	/rpt_family="Alu"	repeat_region	48885. .49110
repeat_region	/rpt_family="SVA"	repeat_region	complement(49395. .49626)
repeat_region	/rpt_family="Alu"	repeat_region	50184. .50437

Query Match	98.8%	Score 2188.6;	DB 9;	Length 142092;
Best Local Similarity	99.5%	Pred. No. 0;		
Matches 2205;	Conservative 0;	Mismatches 10;	Indels 1;	Gaps 1;

1	ATCTCTGTTTCTTTCAGCCTGAAACAAAGAGATGAGTTCAGGCCATTCAGAGG	60
25497	ATCTCTGTTTCTTTCAGCCTGAAACAAAGAGATGAGTTCAGGCCATTCAGAGG	25438
61	ACGAGAGAGATGCTCAAGCTGGAGATATGAGCCGCTTCTAGGCTGCACTGATATC	120
25437	ACGAGAGAGATGCTCAAGCTGGAGATATGAGCCGCTTCTAGGCTGCACTGATATC	25378
121	ACCATCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC	180
25377	ACCATCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC	25318
181	TCCACGAGGCGCTGCCCAAAACCAAGAGCCCAACCAAGAGCTTGGGCGCAGAG	240
25317	TCCACGAGGCGCTGCCCAAAACCAAGAGCCCAACCAAGAGCTTGGGCGCAGAG	25258
241	ACAAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCACTGATAGA	300
25257	ACAAAGAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCACTGATAGA	25198
301	GAGCAGGCTACTTACAGTGGCCCAAGAGCGCCCTACGCCCACTGATGTTTCCGCC	360
25197	GAGCAGGCTACTTACAGTGGCCCAAGAGCGCCCTACGCCCACTGATGTTTCCGCC	25138
361	TGAAACCATCAGCGCCGCTCAAAAGCTTCAAGTGTCAAGGATAGACCAACCAAA	420
25137	TGAAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACCAACCAAA	25078
421	GCTTAAAGCTGTAGTTCTGAGGACCAAGAAAGTTTGAATGCTCTCAGAGGATG	480
25077	GCTTAAAGCTGTAGTTCTGAGGACCAAGAAAGTTTGAATGCTCTCAGAGGATG	25018
481	GGGCTTGATGAGAGACCAAGATATCTCAAGTGAAGAGAAACTGTGATTTAAC	540
25017	GGGCTTGATGAGAGACCAAGATATCTCAAGTGAAGAGAAACTGTGATTTAAC	24958
541	TGACGATATGACAGAGATCCCGAAATCAGCTCAAGAAAGCTTGGAAACATACAA	600
24957	TGACGATATGACAGAGATCCCGAAATCAGCTCAAGAAAGCTTGGAAACATACAA	24898
601	CCAACATACACATACCTCAAGATATCAATGATCTTATTTGGGCTTGGAAACAGT	660
24897	CCAACATACACATACCTCAAGATATCAATGATCTTATTTGGGCTTGGAAACAGT	24838
661	CTGTCTTCACTTGAACAGAGGACATGATTTGCCCTGTGACCCACCCAGCTTCCCTTG	720
24837	CTGTCTTCACTTGAACAGAGGACATGATTTGCCCTGTGACCCACCCAGCTTCCCTTG	24778
721	CTTGAAGCAACCTTCCCAAAATTTCTTGAAGTTCAATGATCTGCAAGACACATGG	780
24777	CTTGAAGCAACCTTCCCAAAATTTCTTGAAGTTCAATGATCTGCAAGACACATGG	24718
781	ACCTAGGCCAGACCTGGCTTGGAGTATATCTTGGCCACTTTCAGATCTTCTTCTT	840
24717	ACCTAGGCCAGACCTGGCTTGGAGTATATCTTGGCCACTTTCAGATCTTCTTCTT	24658

QY	841	AGTGTGGGACCTTTTCTCACTTCAACCTGGATATACCCGGAGAACTTTGGGACAG	900
Db	24657	AGTGTGGGAGCTTTTCTCACTTCAACCTGGATATACCCGGAGAACTTTGGGACAG	24598
QY	901	GTGAAGAAAGTAGAGTTGTGTGCTACCAAGATGCTGTGGAGAATGCCAGGCTGAC	960
Db	24597	GTGAAGAAAGTAGAGTTGTGTGCTACCAAGATGCTGTGGAGAATGCCAGGCTGAC	24538
QY	961	AGGCCAGGCTTAGCTGAGCAGATGTATCACTGGCCCAACTTACTTTGAGCAAGGCTG	1020
Db	24537	AGGCCAGGCTTAGCTGAGCAGATGTATCACTGGCCCAACTTACTTTGAGCAAGGCTG	24478
QY	1021	CTGACCCCAAAACCAATGAGTGGGAGCTGACCTGTGATGACAAATGAACATTCGCCATAC	1080
Db	24477	CTGACCCCAAAACCAATGAGTGGGAGCTGACCTGTGATGACAAATGAACATTCGCCATAC	24418
QY	1081	TATTTAGGGTGTATCCCAAGCACTACAGAAAGGGTGGCAGAACTGCTCACTCTAGG	1140
Db	24417	TATTTAGGGTGTATCCCAAGCACTACAGAAAGGGTGGCAGAACTGCTCACTCTAGG	24358
QY	1141	AACCTGTAGATGTGAGGTTGAGGGGTGCAGCGCCCTTAGGTCAATTTCTCACTGCTG	1200
Db	24357	AACCTGTAGATGTGAGGTTGAGGGGTGCAGCGCCCTTAGGTCAATTTCTCACTGCTG	24298
QY	1201	GGAACTTCACCAAAATACTTCTGCTCTCTTGGGGGTACGCCAAAGCTGTACAAAAATCA	1260
Db	24297	GGAACTTCACCAAAATACTTCTGCTCTCTTGGGGGTACGCCAAAGCTGTACAAAAATCA	24238
QY	1261	GATATTTTCCCTTATTCAGATTTCTGAGACCTGTCAACCATTTTAAACCCCACTT	1320
Db	24237	GATATTTTCCCTTATTCAGATTTCTGAGACCTGTCAACCATTTTAAACCCCACTT	24178
QY	1321	CAGCCCCAATCAGTGGGAGGAAAGTGTAACTTCCCTTTTGTGATTTCTCAAGCAGTTACT	1380
Db	24177	CAGCCCCAATCAGTGGGAGGAAAGTGTAACTTCCCTTTTGTGATTTCTCAAGCAGTTACT	24118
QY	1381	TTCAAGGTCAGAAACAGCAGCTATTATGATTGAACCTTAAAGGCAACAATTTCAAT	1440
Db	24117	TTCAAGGTCAGAAACAGCAGCTATTATGATTGAACCTTAAAGGCAACAATTTCAAT	24059
QY	1441	CTTGCTTCAGGCTTACGATGACAGAACTTGGGAAACATGTGGCTGTCTCAGCAAAAGGATG	1500
Db	24058	CTTGCTTCAGGCTTACGATGACAGAACTTGGGAAACATGTGGCTGTCTCAGCAAAAGGATG	23999
QY	1501	TTCAATATTTAAGAACTCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAGAGAGTCA	1560
Db	23998	TTCAATATTTAAGAACTCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAGAGAGTCA	23939
QY	1561	GGAGTTTGAGACCAACCTGGCCCAACATGATGAACCCCACTCTTACCCAAAAAATACA	1620
Db	23938	GGAGTTTGAGACCAACCTGGCCCAACATGATGAACCCCACTCTTACCCAAAAAATACA	23879
QY	1621	ATCAGGCTGGCCGTGTGTGTGCTGTGATGTCCTCAAGCAGAGAGTTGAGGGAGAAATTGC	1680
Db	23878	ATCAGGCTGGCCGTGTGTGTGCTGTGATGTCCTCAAGCAGAGAGTTGAGGGAGAAATTGC	23819
QY	1681	TTGAACCCAGAGGTTGTGTGTGTGTCAGTGAATTTAGCAACTGCACTCAGCTGGCGAC	1740
Db	23818	TTGAACCCAGAGGTTGTGTGTGTGTCAGTGAATTTAGCAACTGCACTCAGCTGGCGAC	23759
QY	1741	GGAGTGAAGTGTCTCAAAAAAAGATGCTGTCAACCTTTGCTCTCTACT	1800
Db	23758	GGAGTGAAGTGTCTCAAAAAAAGATGCTGTCTCAACCTTTGCTCTCTACT	23699
QY	1801	GCAACATTTTGTGTATTTGAATGAAGGTACCTTCCATCTTAATGCTGTATTAATCTTCAAT	1860
Db	23698	GCAACATTTTGTGTATTTGAATGAAGGTACCTTCCATCTTAATGCTGTATTAATCTTCAAT	23639
QY	1861	TCTCACTAGGATGAACACATTTCTTAACCTGCTCTTAAATGGGATGCTTGGCCAGCA	1920
Db	23638	TCTCACTAGGATGAACACATTTCTTAACCTGCTCTTAAATGGGATGCTTGGCCAGCA	23579
QY	1921	GGTCTCTCACTGTGTGTACACAGCAGAGACATGATTCAGTCAAGCATACAGCTTGC	1980

Db	23578	GGTCTCACTGTGTGTATACACGACGAGACATGATCCAGTCACAGCCATTACAGCTGTCC	23519
Qy	1981	ACACTGAAGAACGTGTCTCTACACACCTGGAATCAATGGTTAGCTTAATAATAAAAT	2040
Db	23518	ACACTGAAGAACGTGTCTCTACACACCTGGAATCAATGGTTAGCTTAATAATAAAAT	23459
Qy	2041	CCGAGACTACTTCAGCTTAAATGCTTTTATTCATPAAAAATGTGAAAGCTAGACTGAA	2100
Db	23458	CCCGAGACTACTTCAGCTTAAATGCTTTTATTCATPAAAAATGTGAAAGCTAGACTGAA	23359
Qy	2101	CCATTGGAAACATTTAATCTCAGACTGTGAGATTGAGAGTCGGGAAACCTTAGTTCTATCTG	2160
Db	23398	CCATTGGAAACATTTAATCTCAGACTGTGAGATTGAGAGTCGGGAAACCTTAGTTCTATCTG	23339
Qy	2161	AATTCAGACAGCCACACTTAATATATCTGCCCAATCTAATAGTTAATAAAATAC	2216
Db	23338	AATTCAGACAGCCACACTTAATATATCTGCCCAATCTAATAGTTAATAAAATAC	23283

LOCUS	AC004228	196080 bp	DNA	linear	PRI 24-FEB-1999
DEFINITION	Homo sapiens Chromosome 11q12.2 PAC clone pDU519013 containing human gene for ferritin heavy chain (FTH), complete sequence.				
ACCESSION	AC004228				
VERSION	AC004228.2	GI:4263838			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 196080)				
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buetter,J., Buneister,R., Card,P., desaillobat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.				
TITLE	HTGS Submissions				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 196080)				
AUTHORS	Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S., Buneister,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.				
TITLE	Direct Submission				
JOURNAL	Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA				
REFERENCE	3 (bases 1 to 196080)				
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buetter,J., Butler,C., Card,P., desaillobat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA				
COMMENT	On Feb 24, 1999 this sequence version replaced gi:291173. IMPORTANT: This submission contains the entire insert of clone pDU519013. pDU519013 comes from the RFL-3 PAC library constructed at the Roswell Park Cancer Institute by the Pletzer de Jong group. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11q12.2 Best's disease region mapped between STS D11S461 and EST AHNAK. This region spans over 1.5 Mbp. MARKER CONFIRMATION: EST, FTH (ferritin heavy chain mRNA), STS6 D11S639 and WI-7524				
	MAPED CLONE OVERLAP: PACs pDU466a1 and pDU756b9.				

FEATURES		Location/Qualifiers
source	1.	.196080
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
repeat_region		complement(2210..2482)
repeat_region		/rpt_family="Alu"
repeat_region		3457..3743
repeat_region		/rpt_family="Alu"
repeat_region		complement(20145..20758)
repeat_region		/rpt_family="Alu"
repeat_region		complement(23194..23599)
repeat_region		/rpt_family="Alu"
repeat_region		complement(27367..27685)
repeat_region		/rpt_family="Alu"
repeat_region		28047..28457
repeat_region		/rpt_family="Alu"
repeat_region		complement(29390..29565)
repeat_region		/rpt_family="Alu"
repeat_region		complement(31076..31255)
repeat_region		/rpt_family="Alu"
repeat_region		complement(34948..35045)
repeat_region		/rpt_family="Alu"
repeat_region		35237..35538
repeat_region		/rpt_family="Alu"
repeat_region		complement(36242..36540)
repeat_region		/rpt_family="Alu"
repeat_region		37238..37522
repeat_region		/rpt_family="Alu"
repeat_region		complement(38555..38650)
repeat_region		/rpt_family="Alu"
repeat_region		complement(40483..41164)
repeat_region		/rpt_family="Alu"
repeat_region		42042..42116
repeat_region		/rpt_family="MIR"
repeat_region		42677..42817
repeat_region		/rpt_family="L1R12"
repeat_region		42856..43157
repeat_region		/rpt_family="Alu"
repeat_region		43148..43380
repeat_region		/rpt_family="L1R12"
repeat_region		complement(43734..43999)
repeat_region		/rpt_family="L1R12"
repeat_region		complement(46998..47169)
repeat_region		/rpt_family="MIR"
repeat_region		complement(48048..48346)
repeat_region		/rpt_family="Alu"
repeat_region		48482..48737
repeat_region		/rpt_family="Alu"
repeat_region		complement(48893..49037)
repeat_region		/rpt_family="L1"
repeat_region		49118..49292
repeat_region		/rpt_family="Alu"
repeat_region		49791..50059
repeat_region		/rpt_family="Alu"
repeat_region		50436..50501
repeat_region		/rpt_family="Alu"
repeat_region		complement(51147..51441)
repeat_region		/rpt_family="Alu"
repeat_region		complement(52700..52979)
repeat_region		/rpt_family="Alu"
repeat_region		53281..53568
repeat_region		/rpt_family="Alu"
repeat_region		54071..54204
repeat_region		/rpt_family="Alu"
repeat_region		54506..54793
repeat_region		/rpt_family="Alu"
repeat_region		54807..55061
repeat_region		/rpt_family="Alu"
repeat_region		complement(64533..64826)
repeat_region		/rpt_family="Alu"
repeat_region		complement(65245..65783)
repeat_region		/rpt_family="Alu"

---

repeat_region		complement(68465..68732)
repeat_region		/rpt_family="Alu"
repeat_region		complement(70191..70476)
repeat_region		/rpt_family="Alu"
repeat_region		70560..70692
repeat_region		/rpt_family="Alu"
repeat_region		72294..72470
repeat_region		/rpt_family="Alu"
repeat_region		complement(72800..73095)
repeat_region		/rpt_family="Alu"
repeat_region		73424..73695
repeat_region		/rpt_family="Alu"
repeat_region		73784..74071
repeat_region		/rpt_family="Alu"
repeat_region		74652..74935
repeat_region		/rpt_family="Alu"
repeat_region		complement(75265..75514)
repeat_region		/rpt_family="Alu"
repeat_region		complement(76086..76353)
repeat_region		/rpt_family="Alu"
repeat_region		complement(78038..78312)
repeat_region		/rpt_family="Alu"
repeat_region		78513..78734
repeat_region		/rpt_family="Alu"
repeat_region		78866..79180
repeat_region		/rpt_family="Alu"
repeat_region		complement(79384..79673)
repeat_region		/rpt_family="Alu"
repeat_region		complement(81398..81690)
repeat_region		/rpt_family="Alu"
repeat_region		82013..82288
repeat_region		/rpt_family="Alu"
repeat_region		complement(83419..83600)
repeat_region		/rpt_family="MIR"
repeat_region		complement(87320..87624)
repeat_region		/rpt_family="Alu"
repeat_region		complement(87710..87987)
repeat_region		/rpt_family="Alu"
repeat_region		complement(88186..88472)
repeat_region		/rpt_family="Alu"
repeat_region		88822..89049
repeat_region		/rpt_family="Alu"
repeat_region		89224..89308
repeat_region		/rpt_family="Alu"
repeat_region		89725..90487
repeat_region		/rpt_family="MER20"
repeat_region		90504..90618
repeat_region		/rpt_family="Alu"
repeat_region		complement(93313..93952)
repeat_region		/rpt_family="Alu"
repeat_region		96938..97517
repeat_region		/rpt_family="Alu"
repeat_region		complement(104199..104514)
repeat_region		/rpt_family="Alu"
repeat_region		complement(109738..109979)
repeat_region		/rpt_family="Alu"
repeat_region		110022..110141
repeat_region		/rpt_family="Alu"
repeat_region		complement(111091..112016)
repeat_region		/rpt_family="Alu"
repeat_region		113210..113419
repeat_region		/rpt_family="L1"
repeat_region		complement(114951..115005)
repeat_region		/rpt_family="MIR"
repeat_region		115685..115791
repeat_region		/rpt_family="MIR"
repeat_region		116870..117005
repeat_region		/rpt_family="MIR"
repeat_region		complement(117877..118172)
repeat_region		/rpt_family="Alu"
repeat_region		118352..118600
repeat_region		/rpt_family="Alu"
repeat_region		118679..118813

```
repeat_region      /rpt_family="Alu"
                   complement(121233. .121520)
                   /rpt_family="Alu"
```

Query Match 98.8%; Score 2188.6; DB 9; Length 196080;

Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2205; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

OY	1	ATCTCTGTTTCTTTCCACCTGAAACAAAGAGAGATGAGTTCCAGGCCCAATCAGAGG	60
Db	191309	ATCTCTGTTTCTTTCCACCTGAAACAAAGAGAGATGAGTTCCAGGCCCAATCAGAGG	191368
OY	61	ACGAGAGAGATGCTCAGCTGGCATTTGCGCGTTCAGAGCTGAGTCTCATGATC	120
Db	191369	ACGAGAGAGATGCTCAGCTGGCATTTGCGCGTTCAGAGCTGAGTCTCATGATC	191428
OY	121	ACCATCTCTCCAGGGCAAACTCAGGACCAAACTACTGTGGCCCAAGGGAAATCCCTTC	180
Db	191429	ACCATCTCTCCAGGGCAAACTCAGGACCAAACTACTGTGGCCCAAGGGAAATCCCTTC	191488
OY	181	TCCAGAGGGGCGTGGCCCAAAAACCAAGGCGCCAAACAGAACGTTATAGGGGCAAGAA	240
Db	191489	TCCAGAGGGGCGTGGCCCAAAAACCAAGGCGCCAAACAGAACGTTATAGGGGCAAGAA	191548
OY	241	ACAAACAAGGCTTGAAAGTTTAAAGCTGTGGAAGCTTCAAGTCTGGCCCCACTGTATCAG	300
Db	191549	ACAAACAAGGCTTGAAAGTTTAAAGCTGTGGAAGCTTCAAGTCTGGCCCCACTGTATCAG	191608
OY	301	GGCCAGGCTTACTACAGTGCCTCCACAGAGCGCCCTCAGCCCTCACTCCCATGTCTTCCCC	360
Db	191609	GGCCAGGCTTACTACAGTGCCTCCACAGAGCGCCCTCAGCCCTCACTCCCATGTCTTCCCC	191668
OY	361	TAGAACCAATCAGAGCGGCTTCAAAAGCTTCAAGTCTCAGGCAATGACACCAAGACAAAA	420
Db	191669	TAGAACCAATCAGAGCGGCTTCAAAAGCTTCAAGTCTCAGGCAATGACACCAAGACAAAA	191728
OY	421	GCTTAAAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATG	480
Db	191729	GCTTAAAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATG	191788
OY	481	GGGCTTGTATGAGACCCCAAGAAATCTCAGAGTGAAGAGAGAAATCTGTGAGTTTAACC	540
Db	191789	GGGCTTGTATGAGACCCCAAGAAATCTCAGAGTGAAGAGAGAAATCTGTGAGTTTAACC	191848
OY	541	TGACGGATATGCGACAGATCCCCGAAAAATCACTCAAGAAAGCTTTGAAACATATCCAA	600
Db	191849	TGACGGATATGCGACAGATCCCCGAAAAATCACTCAAGAAAGCTTTGAAACATATCCAA	191908
OY	601	CCAAACATACACACTACACTCAAAAGATCAGATGGATCCTTATTTGGGCTTGAAACAGGT	660
Db	191909	CCAAACATACACACTACACTCAAAAGATCAGATGGATCCTTATTTGGGCTTGAAACAGGT	191968
OY	661	CTGTCTCTCACTGAAACAGGGGCACTGCATTTGCCCTTGCCCCACCTTCCCTTG	720
Db	191969	CTGTCTCTCACTGAAACAGGGGCACTGCATTTGCCCTTGCCCCACCTTCCCTTG	192028
OY	721	CTGTAGGCTTACCTTCCCTCCCAATTTTCTTGGGTTTCATATCAGTCCAGAGCACTGG	780
Db	192029	CTGTAGGCTTACCTTCCCTCCCAATTTTCTTGGGTTTCATATCAGTCCAGAGCACTGG	192088
OY	781	ACCTACGCCACAGCACTGGCTTGGGGTATATTACTTGGCCACTTTCACAGGATCTTAGGGA	840
Db	192089	ACCTACGCCACAGCACTGGCTTGGGGTATATTACTTGGCCACTTTCACAGGATCTTAGGGA	192148
OY	841	AGTGTTCGGGACCTTTTCTCACTTCACTCCCTGTATACCCGGAAAGCTTCTTGGGACAG	900
Db	192149	AGTGTTCGGGACCTTTTCTCACTTCACTCCCTGTATACCCGGAAAGCTTCTTGGGACAG	192208
OY	901	GTGAAGGAAGATGAGGTTGTGCTGACCAAAATGCTGCTGAGAAACATCTCCACAGGGCTGAC	960
Db	192209	GTGAAGGAAGATGAGGTTGTGCTGACCAAAATGCTGCTGAGAAACATCTCCACAGGGCTGAC	192268
OY	961	AGCCAGAGCTTAGCTGAGCAGATGTTATCATCTGGCCCCCAATTACTTTGACGAAGGATG	1020

Dd	192269	AGCCGAGCTTAACTGAGCGAATGTATTATCATCTGGCCCACTTACTTTAGCAGAAGGTGG	1923288
Oy	1021	CTGACCCAAAACCATGAGGTGGCAGTCAGCTGATGACAGATGAACACTTCCCCTTAAC	1923288
Dd	192329	CTGACCAAAACCATGAGGTGGCACTGACGTGATGACAGATGAACACTTCCCCTTAAC	1923288
Oy	1081	TATTTAGGGTAGTACCCCAAGCATACAGAAAAGGTGGCAGGAACCTGCCTACTTG	1140
Dd	192389	TATTTAGGGTAGTACCCCAAGCATACAGAAAAGGTGGCAGGAACCTGCCTACTTG	192448
Oy	1141	AACGTGTGATGTGTAGGTGTGAGGGGTGCACGGCCCTTAGGTCATTTCTCATCTGCTG	1200
Dd	192449	AACGTGTGATGTGTAGGTGTGAGGGGTGCACGGCCCTTAGGTCATTTCTCATCTGCTG	192508
Oy	1201	GGAACCTCACCAAATACTCTTCTTGCTTCCCTTGGGTGACGCCAAGCTGTCAAAAATCA	1260
Dd	192509	GGAACCTCACCAAATACTCTTCTTGCTTCCCTTGGGTGACGCCAAGCTGTCAAAAATCA	192568
Oy	1261	GATATTTCCCTTATTTCCAGATTTCCCTGGACACTGTACCCAATTATTAACCCCCACTT	1320
Dd	192569	GATATTTCCCTTATTTCCAGATTTCCCTGGACACTGTACCCAATTATTAACCCCCACTT	192628
Oy	1321	CAGCCCCAATCACGTGGGAGGAAGTGTAATCCCTTTCTGTGATTCCTAACAGATTACT	1380
Dd	192629	CAGCCCCAATCACGTGGGAGGAAGTGTAATCCCTTTCTGTGATTCCTAACAGATTACT	192688
Oy	1381	TTACAGGGGTACAGAACGCGAGCTATTATGTAATTGAACCTTAAAAAGGCAACAATTCAANT	1440
Dd	192689	TTACAC-GGTACAGAACGCGAGCTATTATGTAATTGAACCTTAAAAAGGCAACAATTCAANT	192747
Oy	1441	CTTGTCTTAGGCTTAAGACAGGAACCTTGGCAACAATCTGTGGCCTGTGTACGCAAGAGATG	1500
Dd	192748	CTTGTCTTAGGCTTAAGACAGGAACCTTGGCAACAATCTGTGGCCTGTGTACGCAAGAGATG	192807
Oy	1501	TTTCATATTTAAGAATCTTGTCTTGGGGCTGGGTGTGAGGCCAAGTGAATCACAGAGACTCA	1560
Dd	192808	TTAAATATTTAAGAACTTGTCTTGGGGCTGGGTGTGAGGCCAAGTGAATCACAGAGACTCA	192867
Oy	1561	GGAAGTTAGACCAACCTGGCCCAATGATGTGAACCCCACTCTACCAAAAAAAAAATPACA	1620
Dd	192868	GGAAGTTAGACCAACCTGGCCCAATGATGTGAACCCCACTCTACCAAAAAAAAAATPACA	192927
Oy	1621	ATCAGCTGGCGGTGTGTGTGTCTGTATGTCCCAACGCAAGAGTTGAGGGAGAAATTGC	1680
Dd	192928	ATCAGCTGGCGGTGTGTGTGTCTGTATGTCCCAACGCAAGAGTTGAGGGAGAAATTGC	192987
Oy	1681	TTGAACCCAGAGAGTGTGTGTGTGCACTGATGAGCACTGCAATCCAGCTGGGCGAC	1740
Dd	192988	TTGAACCCAGAGAGTGTGTGTGTGCACTGATGAGCACTGCAATCCAGCTGGGCGAC	193047
Oy	1741	GGAGGGAACACTGTCTCAAAAAAAAAAAAAAAAAAGAAATCGCTCAACCTTGGCCCTCTACT	1800
Dd	193048	GGAGGGAACACTGTCTCAAAAAAAAAAAAAAAAAAGAAATCGCTCAACCTTGGCCCTCTACT	193107
Oy	1801	GCAACATTTTGGTATTTGAAATGAAGAGACCTTCCATACTTAATGTCTTAAATACCTTCAAT	1860
Dd	193108	GCAACATTTTGGTATTTGAAATGAAGAGACCTTCCATACTTAATGTCTTAAATACCTTCAAT	193167
Oy	1861	TCTCATAGGAGTGAAGCAATTTCTTAACCTGCTTCTTAATGGGGAATGCTTGGCCGACA	1920
Dd	193168	TCTCATAGGAGTGAAGCAATTTCTTAACCTGCTTCTTAATGGGGAATGCTTGGCCGACA	193227
Oy	1921	GGTCTCTACCTGTGTGTACACGAGAGAGACATGATCCAGTACAGCCCAATACAGCTGTCC	1980
Dd	193228	GGTCTCTACCTGTGTGTACACGAGAGAGACATGATCCAGTACAGCCCAATACAGCTGTCC	193287
Oy	1981	ACACTGAAGAAGTGTCTCTACACAAGCTGTGAATCAATGGTATGCTTAATAGATAAAAT	2040
Dd	193288	ACACTGAAGAAGTGTCTCTCTACACAAGCTGTGAATCAATGGTATGCTTAATAGATAAAAT	193347
Oy	2041	CCCAAGTACTTCAAGCTTTTAATGCTTTTATTCATATAAAACGTGGAAGCTAGACTGA	2100

Db 19348 CCCAGACTACTTACGCTTTAATGCTTTTATATCAATAAACTGTAAAGCTAGCTGAA 193407  
Qy 2101 CCATTGGAACCTTTAACTCAGACTCTTGGAATTCAGAGTCGGGAACCTTAGTTCTATCTG 2160  
Db 193408 CCATTGGAACCTTTAACTCAGACTCTTGGAATTCAGAGTCGGGAACCTTAGTTCTATCTG 193467  
Qy 2161 AATCCAGACAGCCACCTTAGTATCTGCCCAACTAATGATTAATTAATAC 2216  
Db 193468 AATCCAGACAGCCACCTTAGTATCTGCCCAACTAATGATTAATTAATAC 193523

RESUT 4  
AC003025

LOCUS AC003025 112309 bp DNA linear PRI 23-JUL-1998  
DEFINITION Human Chromosome 1p12.2 PAC clone pD466a11, complete sequence.  
ACCESSION AC003025  
VERSION AC003025.1 GI:3337308  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindaes; Homo.  
REFERENCE 1 (bases 1 to 112309)  
AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,  
Buetner,J., Bumeister,R., Card,P., deSallboat,F., Dunn,J.,  
English,C., Ehrhridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,  
Grant,O., Hahner,J., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,  
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,  
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.  
HTGS Submission  
Unpublished  
2 (bases 1 to 112309)  
REFERENCE 2 (bases 1 to 112309)  
AUTHORS Evans,G.A., Athanasiou,M., Bradbury,P., Brignac,S., Bumeister,R.,  
Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M.,  
Gotway,G., Grant,O., Hahner,J., Harris,J., Hinson,S.,  
Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P.,  
Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R.,  
Syed,M., Valenzuela,D., Ward,T. and Wilson,R.  
Direct Submission  
Submitted (21-OCT-1997) Genome Science and Technology Center,  
University of Texas Southwestern Medical Center at Dallas, 5323  
Harry Hines Blvd, Dallas, TX 75235-8591, USA  
3 (bases 1 to 112309)  
REFERENCE 3 (bases 1 to 112309)  
AUTHORS Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,  
Buetner,J., Bumeister,R., Card,P., deSallboat,F., Dunn,J.,  
English,C., Ehrhridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,  
Grant,O., Hahner,J., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,  
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,  
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.  
Direct Submission  
Submitted (23-JUL-1998) Genome Science and Technology Center,  
University of Texas Southwestern Medical Center, 5323 Harry Hines  
Blvd, Dallas, TX 75235-8591, USA  
On Jul 23, 1998 this sequence version replaced gi:2554967.  
pD466a11. This submission contains the entire insert of clone  
pD466a11. pD466a11 comes from a PAC library constructed at the  
Roswell Park Cancer Institute by the Plier de Jong group. This  
clone has been finished according to strict quality criteria and  
as compressions and repetitive elements. The expected phred/Phrap  
calculated errors/10kb is 0.18. In addition, this sequence has  
been finished such that 99.9% of consensus base calls consist of  
either double-stranded coverage or 2 types of labeling chemistry on  
one strand.  
Further information regarding the map of this region or  
annotation of pD466a11 can be found at  
http://gestec.smmcd.edu/chromosome.htm.  
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome  
11p12.2 Best's disease region mapped between STS D11S461 and EST  
AHNAK. This region spans over 1.5 Mbp.  
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA)  
MAPPED CLONE OVERLAP: HTGS submitted PAC clones pD519013 and  
pD475639.

FEATURES  
Source  
Location/Qualifiers  
1..112309  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
repeat\_region  
/rpt\_family="Alu"  
complement(4341..4656)  
repeat\_region  
/rpt\_family="Alu"  
complement(9880..10120)  
repeat\_region  
/rpt\_family="Alu"  
10163..10282  
/rpt\_family="Alu"  
complement(11232..13351)  
repeat\_region  
/rpt\_family="Alu"  
13351..13560  
/rpt\_family="L1"  
complement(15092..15146)  
repeat\_region  
/rpt\_family="MIR"  
15826..15932  
/rpt\_family="MIR"  
17011..17146  
/rpt\_family="MIR"  
complement(18018..18313)  
repeat\_region  
/rpt\_family="Alu"  
18493..18741  
/rpt\_family="Alu"  
18820..18954  
/rpt\_family="Alu"  
complement(21374..22805)  
repeat\_region  
/rpt\_family="Alu"  
22805..23212  
/rpt\_family="Alu"  
23499..23752  
/rpt\_family="Alu"  
24118..24472  
/rpt\_family="Alu"  
25005..25113  
/rpt\_family="MIR"  
31160..31442  
/rpt\_family="Alu"  
complement(32651..32947)  
repeat\_region  
/rpt\_family="Alu"  
complement(35970..36025)  
repeat\_region  
/rpt\_family="MIR"  
38958..39145  
/rpt\_family="Alu"  
44878..45480  
/rpt\_family="Alu"  
49769..50071  
/rpt\_family="Alu"  
complement(50172..50902)  
repeat\_region  
/rpt\_family="Alu"  
50902..50982  
/rpt\_family="MIR"  
51038..51659  
/rpt\_family="Alu"  
54494..54769  
/rpt\_family="Alu"  
55529..55623  
/rpt\_family="THE1"  
complement(55691..55967)  
repeat\_region  
/rpt\_family="Alu"  
complement(56049..56257)  
repeat\_region  
/rpt\_family="Alu"  
56721..57100  
/rpt\_family="THE1"  
58402..58680  
/rpt\_family="Alu"  
complement(58748..59345)  
repeat\_region  
/rpt\_family="Alu"  
59678..59904  
/rpt\_family="Alu"  
61149..61424  
/rpt\_family="Alu"

repeat\_region complement (61930, .62242)  
/rpt\_family="Alu"  
62376, .62710  
repeat\_region /rpt\_family="MLT1"  
62915, .63118  
repeat\_region /rpt\_family="Alu"  
complement (63349, .63455)  
/rpt\_family="Alu"  
64339, .64614  
repeat\_region /rpt\_family="Alu"  
complement (64707, .64984)  
/rpt\_family="Alu"  
repeat\_region complement (65596, .65811)  
/rpt\_family="Alu"  
complement (66500, .66753)  
/rpt\_family="Alu"  
repeat\_region /rpt\_family="Alu"  
67311, .67542  
repeat\_region /rpt\_family="Alu"  
67835, .68119  
repeat\_region /rpt\_family="Alu"  
68860, .69150  
repeat\_region /rpt\_family="Alu"  
69164, .69272  
repeat\_region /rpt\_family="Alu"  
69354, .70245  
repeat\_region /rpt\_family="Alu"  
70655, .70939  
repeat\_region /rpt\_family="Alu"  
70933, .71075  
repeat\_region /rpt\_family="L1"  
71119, .71395  
repeat\_region /rpt\_family="Alu"  
71404, .71530  
repeat\_region /rpt\_family="LTRS"  
71536, .71637  
repeat\_region /rpt\_family="LTRS"  
72720, .72948  
repeat\_region /rpt\_family="L1"  
complement (73488, .73633)  
/rpt\_family="Alu"  
repeat\_region complement (73925, .74200)  
/rpt\_family="Alu"  
75721, .75955  
repeat\_region /rpt\_family="Alu"  
77185, .77449  
repeat\_region /rpt\_family="Alu"  
78093, .78222  
repeat\_region /rpt\_family="L1"  
78227, .78294  
repeat\_region /rpt\_family="L1"  
complement (80458, .80707)  
/rpt\_family="Alu"  
repeat\_region complement (81544, .81974)  
/rpt\_family="Alu"  
82055, .82305  
repeat\_region /rpt\_family="Alu"  
82514, .83096  
repeat\_region /rpt\_family="Alu"  
complement (83685, .83940)  
/rpt\_family="Alu"  
85336, .85850  
repeat\_region /rpt\_family="Alu"  
complement (86429, .86483)  
/rpt\_family="MIR"  
86902, .87183  
repeat\_region /rpt\_family="Alu"  
87646, .87792  
Query Match 98.7%; Score 2187; DB 9; Length 112309;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
QY 1 ATCTCGTTTCTTTCAGCGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGG 60

|||||  
Db 91454 ATCTCGTTTCTTTCAGCGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGG 91513  
QY 61 ACAGAGAGATGCTCAAGCTGGCATCTTGGCCGCTTCTTAGGCTGCAATGCCATGATC 120  
Db 91514 ACAGAGAGATGCTCAAGCTGGCATCTTGGCCGCTTCTTAGGCTGCAATGCCATGATC 91573  
QY 121 ACCATCCTCCAGAGGAAACTCAAGGACCAATCACTGAGGCCCAAGAGGAATCCCTTC 180  
Db 91574 ACCATCCTCCAGAGGCAAACTCAAGGACCAATCACTGAGGCCCAAGAGGAATCCCTTC 91633  
QY 181 TCACAGAGGCGCTGCCCAAAAACCAAGGACCAACCAAGAGCTTAGGCGCCAGAGAG 240  
Db 91634 TCACAGAGGCGCTGCCCAAAAACCAAGGACCAACCAAGAGCTTAGGCGCCAGAGAG 91693  
QY 241 ACAACAAAGGCTGGAAGCTTAAAGCTGTGGAAGCGCTTCAAGTGTGCCCTGATATAGA 300  
Db 91694 ACAACAAAGGCTGGAAGCTTAAAGCTGTGGAAGCGCTTCAAGTGTGCCCTGATATAGA 91753  
QY 301 GGCAGAGCTATCAAGTGGCCCAAGAGCGCGCTCAAGCGCCCACTCCCATGTTCTGCCCC 360  
Db 91754 GGCAGAGCTATCAAGTGGCCCAAGAGCGCGCTCAAGCGCCCACTCCCATGTTCTGCCCC 91813  
QY 361 TAGAACATCAAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAA 420  
Db 91814 TAGAACATCAAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAA 91873  
QY 421 GCTTAAAGACTGTAGTCTTGGGCGCAAGAAAGTTTGAATGCTCTCAAGAGCATG 480  
Db 91874 GCTTAAAGACTGTAGTCTTGGGCGCAAGAAAGTTTGAATGCTCTCAAGAGCATG 91933  
QY 481 GGGCCTTGATGAGCAACCAAGATATCTCAAGTGAAGGAAACGTGGAAGTTAAC 540  
Db 91934 GGGCCTTGATGAGCAACCAAGATATCTCAAGTGAAGGAAACGTGGAAGTTAAC 91993  
QY 541 TGAAGATATGCGAGAGATCCCGAAATCACTCAAGAAACCTTGGAAACAATCAACAA 600  
Db 91994 TGAAGATATGCGAGAGATCCCGAAATCACTCAAGAAACCTTGGAAACAATCAACAA 92053  
QY 601 CCAACATACACTACACTCACTCAAGATCACTGATCTTATTTGGGCTTGGAAAAACAGT 660  
Db 92054 CCAACATACACTACACTCACTCAAGATCACTGATCTTATTTGGGCTTGGAAAAACAGT 92113  
QY 661 CTGTCTCCACCTGAACCAAGGCGCACTGATGCGCGCCGACCCGACTTCCCTTG 720  
Db 92114 CTGTCTCCACCTGAACCAAGGCGCACTGATGCGCGCCGACCCGACTTCCCTTG 92173  
QY 721 CTCTAGACCTACCTTCTCTCACAATTTCTTAGGTTCCATCACTGCGAGAGCACTGG 780  
Db 92174 CTCTAGACCTACCTTCTCTCACAATTTCTTAGGTTCCATCACTGCGAGAGCACTGG 92233  
QY 781 ACCTAGCGCCAGCAGCTGGGCTGGGGTATATCTTGGCCACTTCAAGGAGCTTAGGA 840  
Db 92234 ACCTAGCGCCAGCAGCTGGGCTGGGGTATATCTTGGCCACTTCAAGGAGCTTAGGA 92293  
QY 841 AGTGTTCGGGACCTTTTCTCACTTCAACCTGTGATACCCCGGAAGACTTCTTGGACAG 900  
Db 92294 AGTGTTCGGGACCTTTTCTCACTTCAACCTGTGATACCCCGGAAGACTTCTTGGACAG 92353  
QY 901 GTGAAGAAAGATGAGTGTGCTGACCAAGATCTGTGGAGAACTGCCCCGAGCTGAC 960  
Db 92354 GTGAAGAAAGATGAGTGTGCTGACCAAGATCTGTGGAGAACTGCCCCGAGCTGAC 92413  
QY 961 AGGCAGGCTTAGCTAGCAGAGATGTTATCACTGAGCCCAACTTACTTTGAGCAAGGTGG 1020  
Db 92414 AGGCAGGCTTAGCTAGCAGAGATGTTATCACTGAGCCCAACTTACTTTGAGCAAGGTGG 92473  
QY 1021 CTGAACCAAAACATGAGGTGGAGTCACTGATAGTGAATGAAACATCTCCCACTAAC 1080  
Db 92474 CTGAACCAAAACATGAGGTGGAGTCACTGATAGTGAATGAAACATCTCCCACTAAC 92533  
QY 1081 TATTTAGGTAGTACCAAGCACTACAGAAAGGTGGAGAACTGCGCTAGG 1140  
|||||



```
Dh 92534 TATTTAGGTAATGATCCCAAGCACTAAGGAAAGGTTGCGAAGAACTGCTCACTCTAGG 92593
Qy 1141 AACTGTAATGATGAGGTTGAGGTTGTCAGGCGCTTAGGCAATTTTCTCAGTCTG 1200
Dh 92594 AACTGTAATGATGAGGTTGAGGTTGTCAGGCGCTTAGGCAATTTTCTCAGTCTG 92653
Qy 1201 GGAACCTCAACAAATATCTTCTGCTCTTGGGGTCAAGCCCAAAAGCTGTCAAAATCA 1260
Dh 92654 GGAACCTCAACAAATATCTTCTGCTCTTGGGGTCAAGCCCAAAAGCTGTCAAAATCA 92713
Qy 1261 GATATTTCCCTTTATTCAGATTTCTCGACACTGTCAACCAATTAATAACCCCACTT 1320
Dh 92714 GATATTTCCCTTTATTCAGATTTCTCGACACTGTCAACCAATTAATAACCCCACTT 92773
Qy 1321 CAGCCCAATCACTGAGGAGGAAAGTTAATTTCTTTTCTGATTCGAAGAGTTACT 1380
Dh 92774 CAGCCCAATCACTGAGGAGGAAAGTTAATTTCTTTTCTGATTCGAAGAGTTACT 92833
Qy 1381 TTACAGGAGTCAAGCAAGCACTTATATGATTTGAAACCTTAAAGGCAATTTCACT 92892
Dh 92834 TTACAGGAGTCAAGCAAGCACTTATATGATTTGAAACCTTAAAGGCAATTTCACT 92892
Qy 1441 CTTCCTTTAGGCTTAAGCAAGCACTTGAACATCTGAGCTTTTCAAGAAAGATG 1500
Dh 92893 CTTCCTTTAGGCTTAAGCAAGCACTTGAACATCTGAGCTTTTCAAGAAAGATG 92952
Qy 1501 TTCAATTTTAAGATCTTTGCTTGGGCTGGGTGAGGCAAGTGAATCAAGAGGCTCA 1560
Dh 92953 TTCAATTTTAAGATCTTTGCTTGGGCTGGGTGAGGCAAGTGAATCAAGAGGCTCA 93012
Qy 1561 GGAATTTGAAGCAACCTGCGCAACATGATGAAACCCCATCTCTTCAACAAAAATACA 1620
Dh 93013 GGAATTTGAAGCAACCTGCGCAACATGATGAAACCCCATCTCTTCAACAAAAATACA 93072
Qy 1621 ATAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680
Dh 93073 ATAGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93132
Qy 1681 TTGAACCAAGAGAGTGTGTTGCAAGTGAATGCACTGCAATCACTGAGGAGTGC 1740
Dh 93133 TTGAACCAAGAGAGTGTGTTGCAAGTGAATGCACTGCAATCACTGAGGAGTGC 93192
Qy 1741 GGAAGTGAAGCTGCTCAAAAAAAGATGCTCAACCTTTGCTGCTGCTGCTGCTGCTG 1800
Dh 93193 GGAAGTGAAGCTGCTCAAAAAAAGATGCTCAACCTTTGCTGCTGCTGCTGCTGCTG 93252
Qy 1801 GCAACATTTGATTTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1860
Dh 93253 GCAACATTTGATTTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 93312
Qy 1861 TCTACCTAGGAGTGAAGCACTTTCTTAACTGCTTCTTAAATGAGTGAATTAATCTTCA 1920
Dh 93313 TCTACCTAGGAGTGAAGCACTTTCTTAACTGCTTCTTAAATGAGTGAATTAATCTTCA 93372
Qy 1921 GGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
Dh 93373 GGTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93432
Qy 1981 ACACTGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
Dh 93433 ACACTGAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93492
Qy 2041 CCGAGACTACTGAGCTTTAAATGCTTTTAACTGCTTTTAACTGCTTTTAACTGCTTTTAA 2100
Dh 93493 CCGAGACTACTGAGCTTTAAATGCTTTTAACTGCTTTTAACTGCTTTTAACTGCTTTTAA 93552
Qy 2101 CCAATTGGAACATTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Dh 93553 CCAATTGGAACATTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 93612
Qy 2161 AATTCAGAGCAAGCACTTATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2216
Dh 93613 AATTCAGAGCAAGCACTTATGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 93668
```

```
RESULT 5
AP006260
LOCUS
DEFINITION
  Homo sapiens genomic DNA, chromosome 11 clone:RP11-61G16, complete
ACCESSION
  AP006260
VERSION
  AP006260.2 GI:11790632
KEYWORDS
  HTG.
SOURCE
  Homo sapiens (human)
  Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
  1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
    Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
    Homo sapiens genomic DNA
    Published Only in Database (2003)
  2 (bases 1 to 163024)
  Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
    Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
    Direct Submission
    Submitted (27-FEB-2003) Masahira Hattori, The Institute of Physical
    and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
    1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
    (E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
    Tel:81-45-503-9111, Fax:81-45-503-9170)
    On Jun 16, 2003 this sequence version replaced gi:28812048.
FEATURES
  source
    1..163024
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q"
    /clone="RP11-61G16"
```

## ORIGIN

```
Query Match 98.7%; Score 2187; DB 9; Length 163024;
  Best Local Similarity 99.5%; Pred. No. 0;
  Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

Qy 1 ATCCCTGTTCTTTCTTCCAGCTGAAACAAAGAGAGATGAGTTCAGGCCAATGAGAGG 60
Dh 124425 ATCTCTGTTCTTCTTCCAGCTGAAACAAAGAGAGATGAGTTCAGGCCAATGAGAGG 124484
Qy 61 ACGAGAGAGATGCTCAGCTGCGATCATTTGGCCGCTTCTTGAAGCTTCAAGTCCATGATC 120
Dh 124485 ACGAGAGAGATGCTCAGCTGCGATCATTTGGCCGCTTCTTGAAGCTTCAAGTCCATGATC 124544
Qy 121 ACCATCTCCCGAGGCAACCTGAAGCAAACTACTGTGGCCCAAGAGGAATCCCTTC 180
Dh 124545 ACCATCTCCCGAGGCAACCTGAAGCAAACTACTGTGGCCCAAGAGGAATCCCTTC 124604
Qy 181 TCCACGAGGCGCTGCCCAAAACCAAGAGCAAGCAAAAGCTTAAAGGGCCAGAGAG 240
Dh 124605 TCCACGAGGCGCTGCCCAAAACCAAGAGCAAGCAAAAGCTTAAAGGGCCAGAGAG 124664
Qy 241 ACAACAGGCGCTGGAACCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCAGA 300
Dh 124665 ACAACAGGCGCTGGAACCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCACTGTATCAGA 124724
Qy 301 GGGCAGGCTACTACAGTGCCTCCACAGCGCCCTCAAGCCCACTCCCATGTTCTTCCCTC 360
Dh 124725 GGGCAGGCTACTACAGTGCCTCCACAGCGCCCTCAAGCCCACTCCCATGTTCTTCCCTC 124784
Qy 361 TAGAACCATGAGGCGCTGAAAGCTTCAAGTGTCAAGGCAATAGACCAAAAGCAAAA 420
Dh 124785 TAGAACCATGAGGCGCTGAAAGCTTCAAGTGTCAAGGCAATAGACCAAAAGCAAAA 124844
Qy 421 GCTTAAAGCTGAGTTCGTGGGGCCAGAAAGTTTGAATGCTTCAAGAGCGCATG 480
```



Db 124845 GCTTAAAGACTGTGAGTCTGGGGCCAGAAAAGTTTGAATTGCTCTCAGAGAGCATG 124904  
Qy 481 GGGCCCTTGAATGAGACCCCGAAGATATCTCAATGAGAGAGAAAACCTGTGAGTTTAAACC 540  
Db 124905 GGGCCCTTGAATGAGACCCCGAAGATATCTCAATGAGAGAGAAAACCTGTGAGTTTAAACC 124964  
Qy 541 TGAAGGATATGAGAGATCCCGAAGATATCTCAATGAGAGAGAAAACCTGTGAGTTTAAACC 600  
Db 124965 TGAAGGATATGAGAGATCCCGAAGATATCTCAATGAGAGAGAAAACCTGTGAGTTTAAACC 125024  
Qy 601 CCAACATACACTACCTCAACATGATCATGATCTTATTTGGGCTTTGAAAAACAGGT 660  
Db 125025 CCAACATACACTACCTCAACATGATCATGATCTTATTTGGGCTTTGAAAAACAGGT 125084  
Qy 661 CTGTCTCCACCTGAAACAGAGGGGACCTGCAATGCGCTGTGCCCCACCCAGCTTCCCTTG 720  
Db 125085 CTGTCTCCACCTGAAACAGAGGGGACCTGCAATGCGCTGTGCCCCACCCAGCTTCCCTTG 125144  
Qy 721 CTCTAGAGCTTACCTTCTCCACAAATTTCTAGAGGTTCCATCACTGCGACAGACACTGG 780  
Db 125145 CTCTAGAGCTTACCTTCTCCACAAATTTCTAGAGGTTCCATCACTGCGACAGACACTGG 125204  
Qy 781 ACCTAGCGCCAGCAGCTGCGTTGGGGTATATATCTTGGCACTTTCACAGGGATCTTAGGA 840  
Db 125205 ACCTAGCGCCAGCAGCTGCGTTGGGGTATATATCTTGGCACTTTCACAGGGATCTTAGGA 125264  
Qy 841 AGGTTCGGGACCTTTTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 900  
Db 125265 AGGTTCGGGACCTTTTCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 125324  
Qy 901 GTGAAGAAAGATAGAGTTGTGTGACAGAAATGCTGTGAGAAATGCGCCAGGGCTGAC 960  
Db 125325 GTGAAGAAAGATAGAGTTGTGTGACAGAAATGCTGTGAGAAATGCGCCAGGGCTGAC 125384  
Qy 961 AGGCCAGGCTTACCTGAGCAGATGTTATCACTGCGCCCACTTACTTTAGCAAGAGGTGG 1020  
Db 125385 AGGCCAGGCTTACCTGAGCAGATGTTATCACTGCGCCCACTTACTTTAGCAAGAGGTGG 125444  
Qy 1021 CTGACCCCAAAACATGAGTGGGAGTCACTGATGACAGATGACATTAACATTTCCCAATAC 1080  
Db 125445 CTGACCCCAAAACATGAGTGGGAGTCACTGATGACAGATGACATTAACATTTCCCAATAC 125504  
Qy 1081 TATTTAGGGTAGTACCAGCACTACAGGAAAGGTGGGAGAACTGACCTCACTCTAGG 1140  
Db 125505 TATTTAGGGTAGTACCAGCACTACAGGAAAGGTGGGAGAACTGACCTCACTCTAGG 125564  
Qy 1141 AACTGTAGATGTGAGTGTGAGGGTGTCCAGCGCCCTTAGTATTTTCTCACTGCTG 1200  
Db 125565 AACTGTAGATGTGAGTGTGAGGGTGTCCAGCGCCCTTAGTATTTTCTCACTGCTG 125624  
Qy 1201 GGAACCTCACCAAAATATCTTCTGCTTCTTGGGGTCAAGCCCAAGCTGTCAAAAATCA 1260  
Db 125625 GGAACCTCACCAAAATATCTTCTGCTTCTTGGGGTCAAGCCCAAGCTGTCAAAAATCA 125684  
Qy 1261 GATATTTCCCTTATTCAGATTTCCTGAGCACTGTCAACCAATTTAAACACCCCACT 1320  
Db 125685 GATATTTCCCTTATTCAGATTTCCTGAGCACTGTCAACCAATTTAAACACCCCACT 125744  
Qy 1321 CAGCCCCAATACAGTGGGAGAGTGAATCTTCTTCTGAGTCTCAAGCAGTTACT 1380  
Db 125745 CAGCCCCAATACAGTGGGAGAGTGAATCTTCTTCTGAGTCTCAAGCAGTTACT 125804  
Qy 1381 TTTCAGCGGTCAGAACGCACTATATGATTGAACCTTTAAAGGAGCAAACTTTCAAT 1440  
Db 125805 TTTCAGCGGTCAGAACGCACTATATGATTGAACCTTTAAAGGAGCAAACTTTCAAT 125863  
Qy 1441 CTTCGCTTACAGGTAAGACAGAACTTGGCAAACTGTGGGCTGTTCAGCAAGAGATG 1500  
Db 125864 CTTCGCTTACAGGTAAGACAGAACTTGGCAAACTGTGGGCTGTTCAGCAAGAGATG 125923  
Qy 1501 TTTCATATTTAAGAACTTGTCTTGGGCTGGGTGTGAGGCAAGTCAATCAAGAGATCA 1560  
Db 125924 TTTCATATTTAAGAACTTGTCTTGGGCTGGGTGTGAGGCAAGTCAAGAGATCA 125983

Qy 1561 GGAGTTTGAAGCAACCTGGGCCAATGATGAAACCCCATCTCTACCAAAAAAATACAA 1620  
Db 125984 GGAGTTTGAAGCAACCTGGGCCAATGATGAAACCCCATCTCTACCAAAAAAATACAA 126043  
Qy 1621 ATCAGCTGGCCGTCTGT 1680  
Db 126044 ATCAGCTGGCCGTCTGT 126103  
Qy 1681 TTGAACCCAGAGAGT 1740  
Db 126104 TTGAACCCAGAGAGT 126163  
Qy 1741 GGAGTGAAGCTGTCTCAAAAAAAGAGATGCTCAACCTTTGCCCCCTCTACT 1800  
Db 126164 GGAGTGAAGCTGTCTCAAAAAAAGAGATGCTCAACCTTTGCCCCCTCTACT 126223  
Qy 1801 GCACATTTTGGATTTGAATGAAGTACCTTCCATTAATGCTTTAATCTTTCAAT 1860  
Db 126224 GCACATTTTGGATTTGAATGAAGTACCTTCCATTAATGCTTTAATCTTTCAAT 126283  
Qy 1861 TCTCACTAGGATGAAGCAATTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCA 1920  
Db 126284 TCTCACTAGGATGAAGCAATTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCA 126343  
Qy 1921 GGTCTCACCTGT 1980  
Db 126344 GGTCTCACCTGT 126403  
Qy 1981 ACACGTGAAGACGTGTCTTCAACACAGCCCTGAATCAATAGTTAGTTAATGAATTAAT 2040  
Db 126404 ACACGTGAAGACGTGTCTTCAACACAGCCCTGAATCAATAGTTAATGAATTAAT 126463  
Qy 2041 CCAGACTACTTCAAGCTTTAATGCTTTTATTCATTAATGCTTTAATGCTTTAATGCT 2100  
Db 126464 CCAGACTACTTCAAGCTTTAATGCTTTTATTCATTAATGCTTTAATGCTTTAATGCT 126523  
Qy 2101 CCATTGGAAACATTTAATCACTGAGCTGTGATTCAGAGTGGGGAAACCTTAGTCTATCTG 2160  
Db 126524 CCATTGGAAACATTTAATCACTGAGCTGTGATTCAGAGTGGGGAAACCTTAGTCTATCTG 126583  
Qy 2161 AATCCAAAGCAGCAGCACTTATGATGATGATGATGATGATGATGATGATGATGATGATG 2216  
Db 126584 AATCCAAAGCAGCAGCACTTATGATGATGATGATGATGATGATGATGATGATGATG 126639  
  
RESULT 6  
AC087451/c 163915 bp DNA linear HTG 04-JULY-2001  
LOCUS AC087451 163915 bp DNA linear HTG 04-JULY-2001  
DEFINITION Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT  
ACCESSION AC087451  
VERSION AC087451.2 GI:14595836  
KEYWORDS HTG; HTGS PHASEL; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 163915)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Baerlein, V., Boguslavsky, L., Bouckhaert, B., Brown, A.,  
Cammarata, J., Campoliano, A., Choquet, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodgson, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
Gardyna, S., Glade, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hague, B., Heath, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Karas, A., LaRocque, K., Lamazero, R., Landers, T.,  
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,  
Margulis, N., Mathew, C., McCarthy, M., McEwan, P., McKernan, K.,

TITLE  
JOURNAL

## COMMENT

McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Rector, R.,  
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,  
Roy, A., Santos, R., Schauer, S., Schnepfack, R., Seaman, S., Severy, P.,  
Sougen, C., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, U.,  
Zembek, L., Zimmer, A. and Zody, M.

Submitted (05-JAN-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 4, 2001 this sequence version replaced gi:12039464.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Center project name: L11863

Center clone name: 810.P\_12

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960721

Consensus quality: 162458 bases at least Q40

Consensus quality: 162173 bases at least Q30

Insert size: 170000; agarose-fp

Quality coverage: 10.5 in Q20 bases; agarose-fp

Quality coverage: 10.9 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 17285: contig of 17285 bp in length  
\* 17286 17385: gap of 100 bp  
\* 17386 19170: contig of 1785 bp in length  
\* 19171 19270: gap of 100 bp  
\* 19271 28600: contig of 9330 bp in length  
\* 28601 28700: gap of 100 bp  
\* 28701 47671: contig of 18971 bp in length  
\* 47672 105063: gap of 100 bp  
\* 105064 105163: contig of 57292 bp in length  
\* 105164 142181: gap of 100 bp  
\* 142182 142282: contig of 37018 bp in length  
\* 142282 163915: contig of 100 bp in length.  
\* Location/Qualifiers  
\* 1. 163915

FEATURES  
Source  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"

misc\_feature  
/clone="RP11-810P12"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 17285

misc\_feature  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left"  
17386..19170

## ORIGIN

Query Match 98.7%; Score 2187; DB 2; Length 163915;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

```

QY      1 ATCTCTGTTTCTTTTCAGCTGAACAAAGAGAGATGATGATTCACGCCATCAGAGG 60
DB      151446 ATCTCTGTTTCTTTTCAGCTGAACAAAGAGAGATGATGATTCACGCCATCAGAGG 60
QY      61 ACAGAGAGATGCTGACGCTGCAATCATTTGGCCGCTTCTAGGCTGCAATCCCATGATC 120
DB      151386 ACAGAGAGATGCTGACGCTGCAATCATTTGGCCGCTTCTAGGCTGCAATCCCATGATC 151387
QY      121 ACCATCTCTCCAGGGCAAACTCAAGAGCAAACTAGTGTGGCCCAAGAGAAATCCCTTC 180
DB      151326 ACCATCTCTCCAGGGCAAACTCAAGAGCAAACTAGTGTGGCCCAAGAGAAATCCCTTC 151267
QY      181 TCACAGAGAGGCTGCTCCCAAAACCAAGAGCACCAACAAAGAACTTGAAGGCGCAGAGAG 240
DB      151266 TCACAGAGAGGCTGCTCCCAAAACCAAGAGCACCAACAAAGAACTTGAAGGCGCAGAGAG 151207
QY      241 ACAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGCTGCCCCCATGTTATCAGA 300
DB      151206 ACAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGCTGCCCCCATGTTATCAGA 151147
QY      301 GGCAGGCTACTACATGCCCCCAAGAGGCGCTTCAAGCTTCCCACTGCTTCTTCCCC 360
DB      151146 GGCAGGCTACTACATGCCCCCAAGAGGCGCTTCAAGCTTCCCACTGCTTCTTCCCC 151087
QY      361 TAGAACCATCAGCGCTTAAAGCTTCAAGTGTCAAGAGCATAGAGACCAAGAGCAAA 420
DB      151086 TAGAACCATCAGCGCTTAAAGCTTCAAGTGTCAAGAGCATAGAGACCAAGAGCAAA 151027
QY      421 GCTTAAAGACTGTGATTCGAGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGGATG 480
DB      151026 GCTTAAAGACTGTGATTCGAGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGGATG 150967
QY      481 GGGCCTTGAATGAGAGACCAAGATATCTCAAGTGAAGAGAAACCTGTGGAGATTAAAC 540
DB      150966 GGGCCTTGAATGAGAGACCAAGATATCTCAAGTGAAGAGAAACCTGTGGAGATTAAAC 150907
QY      541 TGACGAGATATGCGAGAGATCCCGAAATCAAGCTCAAAAGAACTTTGGAAACATCAAA 600
DB      150906 TGACGAGATATGCGAGAGATCCCGAAATCAAGCTCAAAAGAACTTTGGAAACATCAAA 150847
QY      601 CCAACATACACATACCTCAAAAGATCAATGATCTTATTTGGCTTGGAAAAACAGGT 660
DB      150846 CCAACATACACATACCTCAAAAGATCAATGATCTTATTTGGCTTGGAAAAACAGGT 150787
QY      661 CTGTCTCTCACTGAAACAGGGGCACTGATGAGCCCTGTGCCCCCAAGCTTCCCTTG 720
DB      150786 CTGTCTCTCACTGAAACAGGGGCACTGATGAGCCCTGTGCCCCCAAGCTTCCCTTG 150727
QY      721 CTCTGAGCTTACCCCTTCTCCCAAACTTCTGAGGTTTCACTGTGCAAGAGCACTGG 780
DB      150726 CTCTGAGCTTACCCCTTCTCCCAAACTTCTGAGGTTTCACTGTGCAAGAGCACTGG 150667
QY      781 ACCTAGCCCAAGCACTGGCTTGGGGTATATCTTGGGCACTTCAAGGATCTTGGGA 840
DB      150667 ACCTAGCCCAAGCACTGGCTTGGGGTATATCTTGGGCACTTCAAGGATCTTGGGA 150667

```

Db	150666	ACTTAGCCAGACCTGCGCTTGGGGATATATCTTGGCCACCTTCACAGGATCTTAGGA	150607
Qy	841	AGTGTTCGGGACCTTTTCTCACTTCAACCTGTATCACCCGGAGAATTCTTGGGACAG	900
Db	150606	AGTGTTGGGAGACTTTTCTCACTTCAACCTGTATCACCCGGAGAATTCTTGGGACAG	150547
Qy	901	GTGAAGGAATGAGGTTGTGCTGACCAATGCTCTGGAGAACCTGCCCGAGGCTGAC	960
Db	150546	GTGAAGGAATGAGGTTGTGCTGACCAATGCTCTGGAGAACCTGCCCGAGGCTGAC	150487
Qy	961	AGGCCAGGCTTAGCTAGCAGATGTATCATCTGAGCCCAACTTACTTTACAGAGGGTGG	1020
Db	150486	AGGCCAGGCTTAGCTAGCAGATGTATCATCTGAGCCCAACTTACTTTAGCAGAGGGTGG	150427
Qy	1021	CTGACCCAAACCATGAGTGGCGAGTCAGCTGGATGACAGATGAACATTCGCCCAATAC	1080
Db	150426	CTGACCCAAACCATGAGTGGCGAGTCAGCTGGATGACAGATGAACATTCGCCCAATAC	150367
Qy	1081	TATTTTAGGTAATGCCCAAGCACTACAGAAAGGGTGGCAGGAATGCTCACTCTTAG	1140
Db	150366	TATTTTAGGTAATGCCCAAGCACTACAGAAAGGGTGGCAGGAATGCTCACTCTTAG	150307
Qy	1141	AACCTGTAATGTTGGAGTTGAGGGGTGCAGGCGCTTAGGCAATTTCTCATCTGCGTG	1200
Db	150306	AACCTGTAATGTTGGAGTTGAGGGGTGCAGGCGCTTAGGCAATTTCTCATCTGCGTG	150247
Qy	1201	GGAACTCTCACAAAATACTTCTTGCTTCCCTGGGGTCAGCCCAAGCTGTCAAAATCA	1260
Db	150246	GGAACTCTCACAAAATACTTCTTGCTTCCCTGGGGTCAGCCCAAGCTGTCAAAATCA	150187
Qy	1261	GATATTTCCCTTTATTCAGATTTCTGGAACACTGTACCCCAATTAAACACCCCACTT	1320
Db	150186	GATATTTCCCTTTATTCAGATTTCTGGAACACTGTACCCCAATTAAACACCCCACTT	150127
Qy	1321	CAGCCCCATCAACGTGGGAGGAAGTAACTTCCCTTTCTGGATTCAGAGAGTACT	1380
Db	150126	CAGCCCCATCAACGTGGGAGGAAGTAACTTCCCTTTCTGGATTCAGAGAGTACT	150067
Qy	1381	TTCAAGGGGTCAACAACAGCAGCTATATGATTTAGAACTTTAAAGGGCAACAATTCAAT	1440
Db	150066	TTCAAC-GGTCAACAACGAGCTATATGATTTAGAACTTTAAAGGGCAACAATTCAAT	150008
Qy	1441	CTTGCTTTAGGCTAAGACAGAACTTGGCAAAACATCTGTGCGCTGTTCAGCAAAAGATG	1500
Db	150007	CTTGCTTTAGGCTAAGACAGAACTTGGCAAAACATCTGTGCGCTGTTCAGCAAAAGATG	149948
Qy	1501	TTCAATATTTAAGAATCTTGTCTTGGGCTGGGTGTGGAAGCAAGTGAATCAAGAGGTCA	1560
Db	149947	TTAAATATTTAAGAATCTTGTCTTGGGCTGGGTGTGGAAGCAAGTGAATCAAGAGGTCA	149888
Qy	1561	GGAGTTTGAAGCAACCTGGCCAACTATATGAAAACCCCATCTCTACCAAAAAATACAA	1620
Db	149887	GGAGTTTGAAGCAACCTGGCCAACTATATGAAAACCCCATCTCTACCAAAAAATACAA	149828
Qy	1621	ATCAGCTGGCCGTGCGTGTGTGCTGTAGTCCCAACGACAGAGGTTAAGGGAGAAATGTC	1680
Db	149827	ATCAGCTGGCCGTGCGTGTGTGCTGTAGTCCCAACGACAGAGGTTAAGGGAGAAATGTC	149768
Qy	1681	TTGAACCCGAGAGGTGTGTGTGCACTGAGATTGAGCACTGCAATCCAGCTGGCGAC	1740
Db	149767	TTGAACCCGAGAGGTGTGTGTGCACTGAGATTGAGCACTGCAATCCAGCTGGCGAC	149708
Qy	1741	GGAGTGAGACTGTCTCAAAAAAAGATCGCTCAACCTTGGAGCTGCTACT	1800
Db	149707	GGAGTGAGACTGTCTCAAAAAAAGATCGCTCAACCTTGGAGCTGCTACT	149648
Qy	1801	GCACATTTTGGTATTTGAAATGAAGTACTTCCATACCTTATGCTTTAATACTTTCAAT	1860
Db	149647	GCACATTTTGGTATTTGAAATGAAGGACTTCCATACCTTATGCTTTAATACTTTCAAT	149588
Qy	1861	TCTCACTAGGAGTGAAGCAATTTCTTAACCTGCTTCTTAATGGGATGCTTTGCCAGCCA	1920
Db	149587	TCTCACTAGGAGTGAAGCAATTTCTTAACCTGCTTCTTAATGGGATGCTTTGCCAGCCA	149528

Qy	1921	GGTCTCTACCTGTGTGTACACACAGCAGSACACTGATTCAGTCAACAGCCATACAGCTGTGC	1980		
Db	149527	GGTCTCTACCTGTGTGTACACACAGCAGSACACTGATTCAGTCAACAGCCATACAGCTGTGC	149468		
Qy	1981	ACACTGAAGAACGTGTCTCTACACAGCCTGAATCAATGSGTTAGCTTAATAGATAAAT	2040		
Db	149467	ACACTGAAGAACGTGTCTCTACACAGCCTGAATCAATGSGTTAGCTTAATAGATAAAT	149444		
Qy	2041	CCCACTACTTCCAGCCTTTAATGCTTTTAATCATAAACCTGTGAACCTAGACTGA	2100		
Db	149407	CCCACTACTTCCAGCCTTTAATGCTTTTAATCATAAACCTGTGAACCTAGACTGA	14933		
Qy	2101	CCATTGGAAACATTTAATCTCAGACTCTGGATTTCAGAGTGGGAAACCTTAGTCTTATC	2160		
Db	149347	CCATTGGAAACATTTAATCTCAGACTCTGGATTTCAGAGTGGGAAACCTTAGTCTTATC	14922		
Qy	2161	AATCCAAGACAGCCACACTTAGTACTGCCCAACTAATAGATTAAATATAC	2216		
Db	149287	AATCCAAGACAGCCACACTTAGTACTGCCCAACTAATAGATTAAATATAC	149232		
RESULT 7	AP003733	166867 bp	DNA	linear	PRI 17-JUN-2003
LOCUS	AP003733				
DEFINITION	Homo sapiens genomic DNA, chromosome 11 clone:RP11-810P12, complete				
ACCESSION	AP003733				
VERSION	AP003733.5				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seop,P., Fujiyama,A., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.				
TITLE	Published Only in Database (2001)				
JOURNAL	2 (bases 1 to 166867)				
REFERENCE	Homo sapiens genomic DNA				
AUTHORS	Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC), 1-7-22 Suenho-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gs.c.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)				
TITLE	Direct Submission				
JOURNAL	On Jun 16, 2003 this sequence version replaced gi:17426128.				
COMMENT	Location/Qualifiers				
FEATURES	1..166867				
Source	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="11"				
	/map="11g"				
	/clone="RP11-810P12"				
ORIGIN					
Query Match	98.7%;	Score 2187;	DB 9;	Length 166867;	
Best Local Similarity	99.5%;	Pred. No. 0;			
Matches 2204;	Conservative	0;	Mismatches 11;	Indels 1;	Gaps 1;
Qy	1	ATCTCTGTTTCTTTCCAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGGAG	60		
Db	12470	ATCTCTGTTTCTTTCCAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGGAG	12529		
Qy	61	ACGAGAGAGATGCTCAGCGTGGCATCATTTGGCCGTTCTTAGGCCCTGACAGTCCCATGATC	120		
Db	12530	ACGAGAGAGATGCTCAGCGTGGCATCATTTGGCCGTTCTTAGGCCCTGACAGTCCCATGATC	12569		
Qy	121	ACGATCTCCACAGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC	180		

Dh 12590 ACCATCCGCCAGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC 12649  
Qy 181 TCCACGAGGGCTCTGCCAAAACCAAGGACGCAAAAGAAAGCTTTGGGGCCGAGAG 240  
Db 12550 TCCACGAGGGCTCTGCCAAAACCAAGGACGCAAAAGAAAGCTTTGGGGCCGAGAG 12709  
Qy 241 ACAACAGGCGCTGGAAGCTTAAAGCTGTGAGAGCCCTTCAAGCTCTGCCACTGTATCAGA 300  
Db 12710 ACAACAGGCGCTGGAAGCTTAAAGCTGTGAGAGCCCTTCAAGCTCTGCCACTGTATCAGA 12769  
Qy 301 GGGCAGGCTTACTCAAGTGGCCCAAGAGCCCTTCAAGCTCTGCCACTGTATCAGA 12769  
Db 12770 GGGCAGGCTTACTCAAGTGGCCCAAGAGCCCTTCAAGCTCTGCCACTGTATCAGA 12829  
Qy 361 TGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATGACACCAAGACAAA 420  
Db 12830 TGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATGACACCAAGACAAA 12889  
Qy 421 GCTTAAAGCTGTGAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGGATG 480  
Db 12890 GCTTAAAGCTGTGAGTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGAGGATG 12949  
Qy 481 GGGCCTTGATGAGACCCAGAAATCTCAAGTGGAGAGAGAAACTGTGAGTTTAAAC 540  
Db 12950 GGGCCTTGATGAGACCCAGAAATCTCAAGTGGAGAGAGAAACTGTGAGTTTAAAC 13009  
Qy 541 TGACGGATATGCGCAGATCCCGAAAATCACTCAAGAAAGCTTTGGAACATGACCA 600  
Db 13010 TGACGGATATGCGCAGATCCCGAAAATCACTCAAGAAAGCTTTGGAACATGACCA 13069  
Qy 601 CCAACATACACACTCACTCAAGATCACTGATGCTTATTTGGGCTTTGGAACATGAC 660  
Db 13070 CCAACATACACACTCACTCAAGATCACTGATGCTTATTTGGGCTTTGGAACATGAC 13129  
Qy 661 CTGTCTCCACCTGGAACAGGGGCACTGATGCTTGTGCCCCCACTTCCCTG 720  
Db 13130 CTGTCTCCACCTGGAACAGGGGCACTGATGCTTGTGCCCCCACTTCCCTG 13189  
Qy 721 CTGTGAGCTTACCTTCTCTCAAAATTTCTTGGGTTTCATCACTGCGACGACACTG 780  
Db 13190 CTGTGAGCTTACCTTCTCTCAAAATTTCTTGGGTTTCATCACTGCGACGACACTG 13249  
Qy 781 ACCTAGGCCGACGCTGGGTTGGGATATATCTTGGCCACTTCAAGGATCTTGAAG 840  
Db 13250 ACCTAGGCCGACGCTGGGTTGGGATATATCTTGGCCACTTCAAGGATCTTGAAG 13309  
Qy 841 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATCACTGGGAAGCTTTGGGACAG 900  
Db 13310 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATCACTGGGAAGCTTTGGGACAG 13369  
Qy 901 GTGAAGGAAGATGAGTGTGCTGACCAAGATGCTGCTGAGAACTGCCCCAGGCTGAC 960  
Db 13370 GTGAAGGAAGATGAGTGTGCTGACCAAGATGCTGCTGAGAACTGCCCCAGGCTGAC 13429  
Qy 961 AGGCGAGGCTTAACTGAGACAGATGTTATCACTGGCCCAACTTCACTTGAAGAGGTTG 1020  
Db 13430 AGGCGAGGCTTAACTGAGACAGATGTTATCACTGGCCCAACTTCACTTGAAGAGGTTG 13489  
Qy 1021 CTGACCCAAAACATGAGTGGGACGTCAGCTGATGATCAAGATGAACTTCCCCATAC 1080  
Db 13490 CTGACCCAAAACATGAGTGGGACGTCAGCTGATGATCAAGATGAACTTCCCCATAC 13549  
Qy 1081 TATTTAGGGTATGATCCCAAGCACTACAGAAAGGGGCAAGAACTGCTCACTCTAGG 1140  
Db 13550 TATTTAGGGTATGATCCCAAGCACTACAGAAAGGGGCAAGAACTGCTCACTCTAGG 13609  
Qy 1141 AACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
Db 13610 AACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 13669  
Qy 1201 GGAACCTCAGCAAAATCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260  
Db 13670 GGAACCTCAGCAAAATCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 13729

Qy 1261 GATATTTCCCTTATATCCAGATTTTCTGAGACCTGTGACCCAAATTAACAACCCCACTT 1320  
Db 13730 GATATTTCCCTTATATCCAGATTTTCTGAGACCTGTGACCCAAATTAACAACCCCACTT 13789  
Qy 1321 CAGCCCAATACGTGGAGAGAGTGTATCTTCCCTTCTGAGATCTCAAGCTTACT 1380  
Db 13790 CAGCCCAATACGTGGAGAGAGTGTATCTTCCCTTCTGAGATCTCAAGCTTACT 13849  
Qy 1381 TTCAAGGATGAGACAGGAGCTTATATGATTTAAACCTTAAAGGCAACATTTCAAT 1440  
Db 13850 TTCAAGGATGAGACAGGAGCTTATATGATTTAAACCTTAAAGGCAACATTTCAAT 13908  
Qy 1441 CTTCCTTCTGAGCTTAAGACAGAACTTGGCAAACTGTGGCTGTTCAAGAAAGATG 1500  
Db 13909 CTTCCTTCTGAGCTTAAGACAGAACTTGGCAAACTGTGGCTGTTCAAGAAAGATG 13968  
Qy 1501 TTCAATTTTAAAGATTTGCTTGGGCTGGTGTGAGAGCAATGATCAGAGAGTCA 1560  
Db 13969 TTCAATTTTAAAGATTTGCTTGGGCTGGTGTGAGAGCAATGATCAGAGAGTCA 14028  
Qy 1561 GAGGTTTGAACCAACCTGGCCCAATGATGAAACCCCATCTTCAACAAAATTAACA 1620  
Db 14029 GAGGTTTGAACCAACCTGGCCCAATGATGAAACCCCATCTTCAACAAAATTAACA 14088  
Qy 1621 ATCAGCTGGCGTGTGGTGTGCTGTATGCTTCCAAAGGAGGTTGAGGGAGAAATTC 1680  
Db 14089 ATCAGCTGGCGTGTGGTGTGCTGTATGCTTCCAAAGGAGGTTGAGGGAGAAATTC 14148  
Qy 1681 TTGAACCCAGAGAGTGTGTGTGAGTGTGAGATTTGAGCACTGCAATCGAGCTGGG 1740  
Db 14149 TTGAACCCAGAGAGTGTGTGTGAGTGTGAGATTTGAGCACTGCAATCGAGCTGGG 14208  
Qy 1741 GAGGTGAGACGTCTCAAAAATTAAGATGCTGTCAACTTTGGCTCTACT 1800  
Db 14209 GAGGTGAGACGTCTCAAAAATTAAGATGCTGTCAACTTTGGCTCTACT 14268  
Qy 1801 GCAACATTTTGTGATTTGAAATGAGATCTTCCATATCTTATGCTGTTAATCTTCAAT 1860  
Db 14269 GCAACATTTTGTGATTTGAAATGAGATCTTCCATATCTTATGCTGTTAATCTTCAAT 14328  
Qy 1861 TCTCATTAGGAGTGAAGACATTTCCATACCTGCTTCTTATGGGAGTCTTGGCAGCA 1920  
Db 14329 TCTCATTAGGAGTGAAGACATTTCCATACCTGCTTCTTATGGGAGTCTTGGCAGCA 14388  
Qy 1921 GGTCTCACCCTGTGTGTATCAACAGAGACACTGATCCAGTCAAGCCTATGCTGCT 1980  
Db 14389 GGTCTCACCCTGTGTGTATCAACAGAGACACTGATCCAGTCAAGCCTATGCTGCT 14448  
Qy 1981 ACACTGAGAGAGTGTCTCTCAACAGCCTGAAATGATGTTAGCTTATGATTAATAT 2040  
Db 14449 ACACTGAGAGAGTGTCTCTCAACAGCCTGAAATGATGTTAGCTTATGATTAATAT 2040  
Qy 2041 CCCAGCTTACTCAGCCTTATATGCTTATATCTTATATCAATTAATGATTAATAT 14508  
Db 14509 CCCAGCTTACTCAGCCTTATATGCTTATATCTTATATCAATTAATGATTAATAT 14568  
Qy 2101 CCATTGAAACATTTAACTCAGACTGATGATGATGATGATGATGATGATGATGAT 2160  
Db 14569 CCATTGAAACATTTAACTCAGACTGATGATGATGATGATGATGATGATGATGAT 14628  
Qy 2161 AATCCAGAGAGCAGCACTTATGATATGCTGCAACCTATATGATGATGATGATGAT 2216  
Db 14629 AATCCAGAGAGCAGCACTTATGATATGCTGCAACCTATATGATGATGATGATGAT 14684

RESULT 8  
AC084857  
LOCUS  
DEFINITION Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT  
ACCESSION AC084857  
VERSION AC084857.1 GI:11276215

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 133683)

JOURNAL Birren, B., Linton, L., Nuebaum, C., and Lander, E.

REFERENCE Homo sapiens chromosome 11, clone CTD-3231N5

AUTHORS 2 (bases 1 to 133683)

Birren, B., Linton, L., Nuebaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Barna, N., Bastien, V., Bede, F., Bogunlavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArliano, K., Dewar, K., Diaz, J. S., Dodge, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Headford, A., Horton, L., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatae, A., Larocque, K., Lamares, R., Landers, T., Lehocsky, J., Levine, R., Lieu, C., Liu, G., Macdonald, P., Margis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T. M., Oliver, J., Peterson, K., Piere, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Sougnez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straube, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tittel, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wymann, D., Ye, W. J., Young, G., Zaimoun, J., Zimmer, A., and Zody, M.

Direct Submission

Submitted (32-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center -----

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

----- Project Information -----

Project name: 3231.N5

Center clone name: 3231.N5

----- Summary Statistics -----

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 126761 bases at least Q40

Consensus quality: 129952 bases at least Q30

Consensus quality: 131351 bases at least Q20

Insert size: 128000; agarose-fp

Insert size: 132483; sum-of-coverage

Quality coverage: 6.1 in Q20 bases; sum-of-coverage

Quality coverage: 5.9 in Q20 bases; sum-of-coverage

----- NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved. -----

1 1459: contig of 1459 bp in length

1460 1559: gap of 100 bp

1560 2520: contig of 961 bp in length

2521 2620: gap of 100 bp

2621 6334: contig of 3714 bp in length

6335 6434: gap of 100 bp

6435 11913: contig of 5479 bp in length

11914 12013: gap of 100 bp

12014 20756: contig of 8743 bp in length

20757 20856: gap of 100 bp

20857 29735: contig of 8879 bp in length

29736 29835: gap of 100 bp

29836 38645: contig of 8810 bp in length

38646 38745: gap of 100 bp

38746 52193: contig of 13448 bp in length

52194 52293: gap of 100 bp

52294 65209: contig of 12916 bp in length

65210 65310: gap of 100 bp

65310 79796: contig of 14467 bp in length

79797 79896: gap of 100 bp

79897 95896: contig of 16000 bp in length

95897 95996: gap of 100 bp

95997 113519: contig of 17423 bp in length

113520 113549: gap of 100 bp

113550 133683: contig of 20164 bp in length.

Location/Qualifiers

1. 133683

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-3231N5"

/clone.lib="CITD Human BAC"

1. 1459

/note="assembly\_fragment"

1560. 2520

/note="assembly\_fragment"

2621. 6334

/note="assembly\_fragment"

6435. 11913

/note="assembly\_fragment"

12014. 20756

/note="assembly\_fragment"

20857. 29735

/note="assembly\_fragment"

29836. 38645

/note="assembly\_fragment"

38746. 52193

/note="assembly\_fragment"

52294. 65209

/note="assembly\_fragment"

65310. 79796

/note="assembly\_fragment"

79897. 95896

/note="assembly\_fragment"

95997. 113519

/note="assembly\_fragment"

113520. 133683

/note="assembly\_fragment"

ORIGIN

Query Match 98.6%; Score 2185.4; DB 2; Length 133683;

Best local Similarity 99.4%; Pred. No. 0;

Matches 2203; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 ATCTCTGTTTCTTTCACCTTGAAACAAAGAGATGATTCAGCCCAATCAGAGG 60

Db 92296 ATCTCTGTTTCTTTCACCTTGAAACAAAGAGATGATTCAGCCCAATCAGAGG 92355

QY 61 ACAGAGAGATGTCACGCTGGCATGATTCGCGCTTCTAGGCGTCAGTCCCATGATC 120

Db 92356 ACAGAGAGATGTCACGCTGGCATGATTCGCGCTTCTAGGCGTCAGTCCCATGATC 92415

QY 121 ACCATCTCCAGGAGCAACTCAGAGCAACTACTGTGCGCCCAAGAGGAATCCCTTC 180

Db 92416 ACCATCTCCAGGAGCAACTCAGAGCAACTACTGTGCGCCCAAGAGGAATCCCTTC 92475

QY 181 TCCACGAGGCGCTGCCCAAAAACACACAGGCGCCCAAGAAAGCTTAAAGGCGCAGAGG 240

Db 92476 TCCACGAGGCGCTGCCCAAAAACACACAGGCGCCCAAGAAAGCTTAAAGGCGCAGAGG 92535

QY 241 ACAACAGAGCCCTGAGAGCTTAAAGCTGTGAGACGCTTCAAGTCTGCCCCCACTGTATACAGA 300  
 DB 92536 ACAACAGAGCCCTGAGAGCTTAAAGCTGTGAGACGCTTCAAGTCTGCCCCCACTGTATACAGA 300  
 QY 301 GCGCAGGCTACTACAGTGGCCCCACAGACGCCCCCTCAGCCCCCATCTCCATGTTCTTCCCCC 360  
 DB 92596 GCGCAGGCTACTACAGTGGCCCCACAGACGCCCCCTCAGCCCCCATCTCCATGTTCTTCCCCC 360  
 QY 361 TAGAACCATCAGGCGCGCTCAAGAGCTTCAAGTGTCAAGAGCAATAGACACCAAGACAAA 420  
 DB 92656 TAGAACCATCAGGCGCGCTCAAGAGCTTCAAGTGTCAAGAGCAATAGACACCAAGACAAA 420  
 QY 421 GCTTAAAGACTGTGAGTCTTGCGGCCAAGAAAGTTTGAATGTCCTCAGAGCGATG 480  
 DB 92716 GCTTAAAGACTGTGAGTCTTGCGGCCAAGAAAGTTTGAATGTCCTCAGAGCGATG 480  
 QY 481 GGGCCTTGTAGAGCAGCCAGAGAGTCTCAAGTGTCAAGAGCAATAGACACCAAGACAAA 540  
 DB 92776 GGGCCTTGTAGAGCAGCCAGAGAGTCTCAAGTGTCAAGAGCAATAGACACCAAGACAAA 540  
 QY 541 TGAACGATATGCAAGATCCCGAATACTCAAGAGCAATAGACACCAAGACAAA 600  
 DB 92836 TGAACGATATGCAAGATCCCGAATACTCAAGAGCAATAGACACCAAGACAAA 600  
 QY 601 CCAACATACACACTACCTCAAGAGTCAAGTGTCTTATTTGGGCTTGGAAACAGAGT 660  
 DB 92896 CCAACATACACACTACCTCAAGAGTCAAGTGTCTTATTTGGGCTTGGAAACAGAGT 660  
 QY 661 CTGTCTTCACCTGACCAAGAGGACATGATTCCTTATTTGGGCTTGGAAACAGAGT 720  
 DB 92956 CTGTCTTCACCTGACCAAGAGGACATGATTCCTTATTTGGGCTTGGAAACAGAGT 720  
 QY 721 CTGTAGACCTACCTTCTCCACATTTCTAGAGGTTCCATCACTGCGACAGACACTGG 780  
 DB 93016 CTGTAGACCTACCTTCTCCACATTTCTAGAGGTTCCATCACTGCGACAGACACTGG 780  
 QY 781 ACCTAGCCGAGAGCTGCTTGGGATATATCTTGGCCACTTCACTGCGACAGACACTGG 93075  
 DB 93076 ACCTAGCCGAGAGCTGCTTGGGATATATCTTGGCCACTTCACTGCGACAGACACTGG 93075  
 QY 841 AGTGTTCGAGACCTTCTTCACTTCACTGCGACAGAGTCTTGGGACAG 900  
 DB 93136 AGTGTTCGAGACCTTCTTCACTTCACTGCGACAGAGTCTTGGGACAG 900  
 QY 901 GTGAAGAGAGTATGAGTGTCTGACAGAAATGCTGTGAGAACTGCCCCAGAGGCTGAC 93255  
 DB 93196 GTGAAGAGAGTATGAGTGTCTGACAGAAATGCTGTGAGAACTGCCCCAGAGGCTGAC 93255  
 QY 961 AGCCAGGCTTAGCTGAGACAGATGTATCACTGCGCCCACTTATTTAGCAAGGCTG 1020  
 DB 93256 AGCCAGGCTTAGCTGAGACAGATGTATCACTGCGCCCACTTATTTAGCAAGGCTG 1020  
 QY 1021 CTGACCCCAAAACATGAGGCTGAGTGTGCTGACAGAAATGCTGTGAGAACTGCCCCAGAGGCTG 93315  
 DB 93316 CTGACCCCAAAACATGAGGCTGAGTGTGCTGACAGAAATGCTGTGAGAACTGCCCCAGAGGCTG 93315  
 QY 1081 TATTTAGGATATGATCCCAAGACCTAGAGAAAGGCTGAGAGAACTGCTCACTCTAGG 1140  
 DB 93376 TATTTAGGATATGATCCCAAGACCTAGAGAAAGGCTGAGAGAACTGCTCACTCTAGG 1140  
 QY 1141 AACTGTAGATGAGGAGTGTGAGGCTGACAGGCGCTTAAAGTCTTCTCACTGCTG 1200  
 DB 93436 AACTGTAGATGAGGAGTGTGAGGCTGACAGGCGCTTAAAGTCTTCTCACTGCTG 1200  
 QY 1201 GGAACCTCAACAAATATCTTCTGCTTGGGCTGAGCCCAAGCTGTCACAAAATCA 1260  
 DB 93496 GGAACCTCAACAAATATCTTCTGCTTGGGCTGAGCCCAAGCTGTCACAAAATCA 1260  
 QY 1261 GATATTTCCCTTATTTCCAGATTTCTGAGACCTGTCAACCAATATATACACCCACTT 1320  
 DB 93556 GATATTTCCCTTATTTCCAGATTTCTGAGACCTGTCAACCAATATATACACCCACTT 1320  
 QY 1321 CAGCCCAATACAGTGGAGAGATGTATCTTCTTGTGATTTCTAAGCACTTACT 1380

DB 93616 CAGCCCAATACAGTGGAGAGATGTATCTTCTTCTGATTTCTCAAGAGTACT 93675  
 QY 1381 TTCAAGGCTGAGACAGCAGACTATTTATGATTTGAATCTTAAAGGCAACATTTGANT 1440  
 DB 93676 TTCAAGGCTGAGACAGCAGACTATTTATGATTTGAATCTTAAAGGCAACATTTGANT 1440  
 QY 1441 CTTCGCTTACAGCTTAAAGCAGAACTTGGCAACATCTGTGCTGTTGAGCAAGAGATG 1500  
 DB 93735 CTTCGCTTACAGGCTTAAAGCAGAACTTGGCAACATCTGTGCTGTTGAGCAAGAGATG 1500  
 QY 1501 TTCAATTTTAAAGATCTTGTCTTGGCTGGTGTGAGAGCAAGTAAATCAAGAGGCTCA 1560  
 DB 93795 TTCAATTTTAAAGATCTTGTCTTGGCTGGTGTGAGAGCAAGTAAATCAAGAGGCTCA 1560  
 QY 1561 GGAATTTGAAGACCAACCTGGCCCAACATGATGAACCCCATCTTACCAAAAAATACAA 1620  
 DB 93855 GGAATTTGAAGACCAACCTGGCCCAACATGATGAACCCCATCTTACCAAAAAATACAA 1620  
 QY 1621 ATCAGCTGCGCGCTGTGTGTGCTGTAGTCCCAAGCAGAGGCTGAGAGAGATTC 1680  
 DB 93915 ATCAGCTGCGCGCTGTGTGTGCTGTAGTCCCAAGCAGAGGCTGAGAGAGATTC 1680  
 QY 1681 TTGAACCCAGAGAGTGTGTGTGTGAGATTTAGCACTGCAATCAGCTGGCGAG 1740  
 DB 93975 TTGAACCCAGAGAGTGTGTGTGTGAGATTTAGCACTGCAATCAGCTGGCGAG 1740  
 QY 1741 GGAAGTGAAGCTGTCTCAAAAAAAGATGCTCAACCTTGGCTCTACT 1800  
 DB 94035 GGAAGTGAAGCTGTCTCAAAAAAAGATGCTCAACCTTGGCTCTACT 1800  
 QY 1801 GCAACATTTGTGATTTGAATGAAGTACTTCCATCTTATGCTGTATATCTTCAAT 1860  
 DB 94095 GCAACATTTGTGATTTGAATGAAGTACTTCCATCTTATGCTGTATATCTTCAAT 1860  
 QY 1861 TCTACTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 DB 94155 TCTACTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920  
 QY 1921 GGTCTCACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 DB 94215 GGTCTCACTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980  
 QY 1981 ACACTGAAGAGCTGTCTTCAACAGCTGATCAAGCTGATCAAGCTGATCAAGCTGAT 1981  
 DB 94275 ACACTGAAGAGCTGTCTTCAACAGCTGATCAAGCTGATCAAGCTGATCAAGCTGAT 1981  
 QY 2041 CCAAGACTTCTTCAAGCTTCTTATGCTTCTTATGATGATGATGATGATGATGATGATGAT 2100  
 DB 94335 CCAAGACTTCTTCAAGCTTCTTATGCTTCTTATGATGATGATGATGATGATGATGATGAT 2100  
 QY 2101 CCATTTGAAACATTTAATCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 2160  
 DB 94395 CCATTTGAAACATTTAATCTGATCTGATGATGATGATGATGATGATGATGATGATGAT 2160  
 QY 2161 AATCAAGACAGCAGACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2216  
 DB 94455 AATCAAGACAGCAGACCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2216  
 RESULT 9  
 AC051664/c  
 LOCUS  
 DEFINITION Homo sapiens chromosome 11 clone RP11-810P12, WORKING DRAFT  
 ACCESSION AC051664  
 VERSION AC051664.2 GI:7637836  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 1 (bases 1 to 16169)



AUTHORS Waterston,R.H.  
TITLE The sequence of Homo sapiens clone  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 160169)  
AUTHORS Waterston,R.H.  
JOURNAL Direct Submision  
Submitted (15-APR-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Apr 23, 2000 this sequence version replaced gi:1574984.

## COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
----- Project Information -----  
Center project name: H\_NH0810P12  
----- Summary Statistics -----  
Sequencing vector: M13, 100%  
Sequencing vector: plasmid: 0%  
Chemistry: Dye-primer ET, 100% of reads  
Chemistry: Dye-terminator Big Dye, 0% of reads  
Assembly program: Phrap, Version 0.990319  
Consensus quality: 147316 bases at least Q40  
Consensus quality: 151042 bases at least Q30  
Consensus quality: 152875 bases at least Q20  
Insert size: 165000; agarose-fp  
Insert size: 157869; sum-of-contigs  
Quality coverage: 3.91 in Q20 bases; sum-of-contigs  
Quality coverage: 4.17 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 24 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1926: contig of 1926 bp in length  
\* 1927 2026: gap of unknown length  
\* 2027 4202: contig of 2176 bp in length  
\* 4203 4302: gap of unknown length  
\* 4303 7163: contig of 2861 bp in length  
\* 7164 7263: gap of unknown length  
\* 7264 9961: contig of 2698 bp in length  
\* 9962 10061: gap of unknown length  
\* 10062 12575: contig of 2514 bp in length  
\* 12576 12675: gap of unknown length  
\* 12676 15545: contig of 2870 bp in length  
\* 15546 15645: gap of unknown length  
\* 15646 20941: contig of 5296 bp in length  
\* 20942 21041: gap of unknown length  
\* 21042 23966: contig of 2925 bp in length  
\* 23967 24066: gap of unknown length  
\* 24067 27231: contig of 3165 bp in length  
\* 27232 27331: gap of unknown length  
\* 27332 30550: contig of 3219 bp in length  
\* 30551 30650: gap of unknown length  
\* 30651 35202: contig of 4552 bp in length  
\* 35203 35302: gap of unknown length  
\* 35303 40301: contig of 4899 bp in length  
\* 40302 40301: gap of unknown length  
\* 40302 45758: contig of 5457 bp in length  
\* 45759 45858: gap of unknown length  
\* 45859 51249: contig of 5391 bp in length  
\* 51250 51349: gap of unknown length  
\* 51350 58267: contig of 6918 bp in length  
\* 58268 58367: gap of unknown length  
\* 58368 66916: contig of 8549 bp in length  
\* 66917 77145: contig of 10129 bp in length  
\* 77146 77245: gap of unknown length

\* 77246 86493: contig of 9248 bp in length  
\* 86494 86593: gap of unknown length  
\* 86594 96688: contig of 10095 bp in length  
\* 96689 96788: gap of unknown length  
\* 96789 107169: contig of 10381 bp in length  
\* 107170 107269: gap of unknown length  
\* 107270 117047: contig of 9778 bp in length  
\* 117048 117147: gap of unknown length  
\* 117148 127120: contig of 9973 bp in length  
\* 127121 127220: gap of unknown length  
\* 127221 142029: contig of 14809 bp in length  
\* 142030 142129: gap of unknown length  
\* 142130 160169: contig of 18040 bp in length.  
\* 160169 Location/Qualifiers  
1. 160169  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/clone="RP11-810P12"  
1. 1926  
/note="assembly\_name:Contig18"  
2027. 4202  
/note="assembly\_name:Contig19"  
4303. 7163  
/note="assembly\_name:Contig20  
clone\_end:"7"  
vector\_side:left"  
7264. 9961  
/note="assembly\_name:Contig21"  
10062. 12575  
/note="assembly\_name:Contig22"  
12676. 15545  
/note="assembly\_name:Contig23"  
15646. 20941  
/note="assembly\_name:Contig24"  
21042. 23966  
/note="assembly\_name:Contig25"  
24067. 27231  
/note="assembly\_name:Contig26"  
27332. 30550  
/note="assembly\_name:Contig27"  
30651. 35202  
/note="assembly\_name:Contig28"  
35303. 40201  
/note="assembly\_name:Contig29"  
40302. 45758  
/note="assembly\_name:Contig30"  
45859. 51249  
/note="assembly\_name:Contig31"  
51350. 58267  
/note="assembly\_name:Contig32"  
58368. 66916  
/note="assembly\_name:Contig33"  
67017. 77145  
/note="assembly\_name:Contig34"  
77246. 86493  
/note="assembly\_name:Contig35"  
86594. 96688  
/note="assembly\_name:Contig36"  
96789. 107169  
/note="assembly\_name:Contig37"  
107270. 117047  
/note="assembly\_name:Contig38"  
117148. 127120  
/note="assembly\_name:Contig39  
clone\_end:SP6  
vector\_side:right"  
127221. 142029  
/note="assembly\_name:Contig40"  
142130. 160169  
/note="assembly\_name:Contig41"

ORIGIN

Query Match	72.6%	Score 1608.4	DB 2	Length 160169
Best Local Similarity	99.1%	Pred. No. 0		
Matches 1626	Conservative 0	Mismatches 13	Indels 1	Gaps 1
QY 577	AAGAACCTTTGGAAACATCAACCAACATCAATCACTACACTCAAGATCAATGATGATC	636		
Db 51139	AAGACCNNTTGGAAACATCAACCAACATCAATCACTACACTCAAGATCAATGATGATC	51080		
QY 637	CTTATTTGGGCGCTTGGAAAAACAGGCTGTCTCTCCACTGAAACAGGGGCACTGATGGCC	696		
Db 51079	CTTATTTGGGCGCTTGGAAAAACAGGCTGTCTCTCCACTGAAACAGGGGCACTGATGGCC	51020		
QY 697	TGTGCCCCCACTCCAGCTTCTCTTGTCTGTAGAGCTAACCTTCCCTCCAAATTTCTTAGGCT	756		
Db 51019	TGTGCCCCCACTCCAGCTTCTCTTGTCTGTAGAGCTAACCTTCCCTCCAAATTTCTTAGGCT	50960		
QY 757	TTCATCACTGACAGAGACACATGAGACCTTACGCCACACTGCTGTGGGTATATACCTTGG	816		
Db 50959	TTCATCACTGACAGAGACACACTGACCTTACGCCACACTGCTGTGGGTATATACCTTGG	50900		
QY 817	CCACCTTCCAGGAGATCTTAGAGGAATGTTTGGGACCTTTTCTCACTTCACTCCGTGATC	876		
Db 50899	CCACCTTCCAGGAGATCTTAGAGGAATGTTTGGGACCTTTTCTCACTTCACTCCGTGATC	50840		
QY 877	ACCCGGAAGACTTCTTGGGACAGGTCAGGTAAGGAAGATGAGGTTTGTGCTGACCAAGATCTG	936		
Db 50839	ACCCGGAAGACTTCTTGGGACAGGTCAGGTAAGGAAGATGAGGTTTGTGCTGACCAAGATCTG	50780		
QY 937	CTGAGAACTGCCCCAGGGCTGACAGGCTCAGGCTTAACTGACAGATGTTATCACTGGC	996		
Db 50779	CTGAGAACTGCCCCAGGGCTGACAGGCTCAGGCTTAACTGACAGATGTTATCACTGGC	50720		
QY 997	CCAATTTACTTTGACAAAGGGTGGCTGACCCAAACCATGAGGGGAGCTGACACTGGATG	1056		
Db 50719	CCAATTTACTTTGACAAAGGGTGGCTGACCCAAACCATGAGGGGAGCTGACACTGGATG	50660		
QY 1057	ACAGATGACACTTCCCCCAATTAATATTTAGGTAAGTACCAGCACTACAGAAAGGGT	1116		
Db 50659	ACAGATGACACTTCCCCCAATTAATATTTAGGTAAGTACCAGCACTACAGAAAGGGT	50600		
QY 1117	GGCAGGAACGTGCTCATCTCTAGGAACCTGTAGATGATGATGATGAGGTTGACAGCCG	1176		
Db 50599	GGCAGGAACGTGCTCATCTCTAGGAACCTGTAGATGATGATGATGAGGTTGACAGCCG	50540		
QY 1177	CTTAGGTCAATTTCTCACTGCTGGGAACTCAACCAATATCTTCTGCTTCTTGGGAT	1236		
Db 50539	CTTAGGTCAATTTCTCACTGCTGGGAACTCAACCAATATCTTCTGCTTCTTGGGAT	50480		
QY 1237	CAGCCCAAGGCTGTCAAAATCATAGATATTTCCCTTTATTCAGATTTCTCGACACTTT	1296		
Db 50479	CAGCCCAAGGCTGTCAAAATCATAGATATTTCCCTTTATTCAGATTTCTCGACACTTT	50420		
QY 1297	CACCCAAATTAATAACACCCCACTTACGCCCAATCAGCTGGGAGGAATGTAACTTCCCT	1356		
Db 50419	CACCCAAATTAATAACACCCCACTTACGCCCAATCAGCTGGGAGGAATGTAACTTCCCT	50360		
QY 1357	TTTTCGATTTCTCAAGCAGTTACTTTTCACTCGGCTCAGAAACGCGACTATTAATGAA	1416		
Db 50359	TTTTCGATTTCTCAAGCAGTTACTTTTCACTCGGCTCAGAAACGCGACTATTAATGAA	50300		
QY 1417	CCTTAATAAGGCAACAATTTCAATCTTGTCTTCTAGCTAAGACAGGAACCTGGCAACAT	1476		
Db 50300	CCTTAATAAGGCAACAATTTCAATCTTGTCTTCTAGCTAAGACAGGAACCTGGCAACAT	50240		
QY 1477	CTGTGGCTGTTCAGCAAAAGATGTGTTATTTAAAGATTTTGTCTTGGCTGGGTGG	1536		
Db 50240	CTGTGGCTGTTCAGCAAAAGATGTGTTATTTAAAGATTTTGTCTTGGCTGGGTGG	50180		
QY 1537	AGGCAATGTAATCAAGAGAGTCAAGAGTTTGAACCAACTGGCCAACTATGTAAC	1596		
Db 50180	AGGCAATGTAATCAAGAGAGTCAAGAGTTTGAACCAACTGGCCAACTATGTAAC	50120		

QY	1597	CCACTCTACCAAAAAAAAAATCAATCAAGTCAGCGCGTGTGTGCGCTGTAGTCCCAAC	1656
Db	50120	CCACTCTACCAAAAAAAAAATCAAAATCAAGTCAGCGCGTGTGTGCTGTAGTCCCAAC	50067
QY	1657	GCAGAGAGTTGAGGGAGAAATGCTGTAACCCAGAGAGGTGGTGGTTCAGTGAATTGAG	1716
Db	50060	GCAGAGAGTTGAGGGAGAAATGCTTGAACCCAGAGAGGTGGTGGTTCAGTGAATTGAG	50001
QY	1717	CAACTGAATCCAGCGCTGGGCGACGAGTGAAGTGTCTCAAAAAAAAAAAAAAGAT	1776
Db	50000	CAACTGAATCCAGCGCTGGGCGACGAGAGTGAAGTGTCTCAAAAAAAAAAAAAAGAT	49941
QY	1777	CGTCTCAACCTTGGCTTCCTCACTGCAACATTTGGTAATTGAAATGAAGATACCTTCCA	1836
Db	49940	CGTCTCAACCTTGGCTTCCTCACTGCAACATTTGGTAATTGAAATGAAGAGCTTCCA	49881
QY	1837	TACTATGTGTTATTAATTTTCATTTCTCACTAGGATGAAGCAATCTCTCAACTGCTTC	1896
Db	49880	TACTATGTGTTATTAATTTTCATTTCTCACTAGGATGAAGCAATTTCTCAACTGCTTC	49821
QY	1897	CTAAAGGGAGATGCTTCGCGACGCGAGTCTCACTGTGTGTATCACACAGAGAGACATGAT	1956
Db	49820	CTAAAGGGAGATGCTTCGCGACGCGAGTCTCACTGTGTGTATCACACAGAGAGACATGAT	49761
QY	1957	CCAGTCACAGCCATACAGCTGTGCACACATGGAAGAAAGTGTCTACACAGCTGAATCAA	2016
Db	49760	CCAGTCACAGCCATACAGCTGTGCACACATGGAAGAAAGTGTCTACACAGCTGAATCAA	49701
QY	2017	ATGGTTAGCTTAATTAAGTAATAAATCCGAGCTACTTCAGCCTTTAATGCTTTAATTCAT	2076
Db	49700	ATGGTTAGCTTAATTAAGTAATAAATCCGAGCTACTTCAGCCTTTAATGCTTTAATTCAT	49641
QY	2077	AAAAACCTGTAAAGCTAGACTGAACCATTTGAAAAATTAACTAGACTGTGAATTCAGA	2136
Db	49640	AAAAACCTGTAAAGCTAGACTGAACCATTTGAAAAATTAACTAGACTGTGAATTCAGA	49581
QY	2137	GTCGGAAACCTTAGTGTCTATCTGAATCCAGAGACGCAACCTTAGTATCTGCCCAA	2196
Db	49580	GTCGGAAACCTTAGTGTCTATCTGAATCCAGAGACGCAACCTTAGTATCTGCCCAA	49521
QY	2197	CTAATAGTGAATTAATAATAC	2216
Db	49520	CTAATAGTGAATTAATAATAC	49501

RESULT	10
AC0004586/c	
LOCUS	
DEFINITION	AC004588 106648 bp DNA linear HTG 13-JUN-2002
ACCESSION	Homo sapiens, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces.
VERSION	AC004588
KEYWORDS	AC004588.2 GI:21406530
SOURCE	HTG; HTGS PHASE1
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 106648)
AUTHORS	Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basil,M., Beutner,J., Bumelster,R., Card,P., desaiIboat,C., Dunn,J., English,C., Ehrbridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,U., Harris,U., Lewis,E., Loo,H., Loo,K.N., McFarland,J., Miller,R., Newton,J., Osborne-Lawrence,S., Schagman,U., Schultz,R.A., Stimson,S., Syed,M. and Ward,T. HTS Submission Unpublished 2 (bases 1 to 106648) Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basil,M., Beutner,J., Bumelster,R., Card,P., desaiIboat,C., Dunn,J., English,C., Ehrbridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,U., Harris,U., Lewis,E., Loo,H., Loo,K.N., McFarland,J., Miller,R., Newton,J., Osborne-Lawrence,S., Schagman,U., Schultz,R.A., Stimson,S., Syed,M. and Ward,T. Direct Submission
TITLE	



## JOURNAL

Submitted (18-APR-1998) Genome Science & Technology Center,  
University of Texas Southwestern Medical Center, 5323 Harry Hines  
Blvd, Dallas, TX 75235-8591, USA

## COMMENT

On Jun 13, 2002 this sequence version replaced gi:3063499.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 19 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
\* 2483: contig of 2483 bp in length  
\* 2484 2583: gap of 100 bp  
\* 2584 4622: contig of 2039 bp in length  
\* 4623 4722: gap of 100 bp  
\* 4723 6725: contig of 2003 bp in length  
\* 6726 6825: gap of 100 bp  
\* 6826 8448: contig of 1623 bp in length  
\* 8449 8548: gap of 100 bp  
\* 8549 11050: contig of 2502 bp in length  
\* 11051 11150: gap of 100 bp  
\* 11151 13554: contig of 2404 bp in length  
\* 13555 13654: gap of 100 bp  
\* 13655 17064: contig of 3410 bp in length  
\* 17065 17164: gap of 100 bp  
\* 17165 20800: contig of 3636 bp in length  
\* 20801 20900: gap of 100 bp  
\* 20901 23967: contig of 3067 bp in length  
\* 23968 24067: gap of 100 bp  
\* 24068 26216: contig of 2149 bp in length  
\* 26217 26316: gap of 100 bp  
\* 26317 31202: contig of 4886 bp in length  
\* 31203 31302: gap of 100 bp  
\* 31303 35714: contig of 4412 bp in length  
\* 35715 35814: gap of 100 bp  
\* 35815 39320: contig of 3506 bp in length  
\* 39321 39420: gap of 100 bp  
\* 39421 47809: contig of 8389 bp in length  
\* 47810 47909: gap of 100 bp  
\* 47910 55372: contig of 7463 bp in length  
\* 55373 55472: gap of 100 bp  
\* 55473 56654: contig of 1182 bp in length  
\* 56655 56754: gap of 100 bp  
\* 56755 65820: contig of 9086 bp in length  
\* 65821 65920: gap of 100 bp  
\* 65921 87612: contig of 21692 bp in length  
\* 87613 87712: gap of 100 bp  
\* 87713 106648: contig of 18936 bp in length.

## FEATURES

## source

1..106648  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
1..104848  
/note="CHROMOSOMAL LOCUS: This PAC clone comes from the  
Chromosome -  
11p12.2 Best's disease region mapped between STS D11S461  
and EST AHNNAK. This region spans over 1.5 Mbp.  
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA)

## misc\_feature

MAPPED CLONE OVERLAP: HTGS phase 1 PAC clone pdu466a11"

repeat\_region  
473..745  
/rpt\_family="Alu"  
complement(1443..1649)  
repeat\_region  
/rpt\_family="MER21"  
complement(2690..2959)  
repeat\_region  
/rpt\_family="Alu"  
3427..3725  
/rpt\_family="Alu"  
complement(5231..5605)  
repeat\_region  
/rpt\_family="LTR3"  
complement(6706..6996)

/rpt\_family="Alu"  
7018..7459  
/rpt\_family="MER21"  
7067..7420  
/rpt\_family="MER39"  
complement(7570..7856)  
repeat\_region  
/rpt\_family="Alu"  
11234..11801  
/rpt\_family="Alu"  
complement(12283..12373)  
repeat\_region  
/rpt\_family="MIR"  
12666..13031  
/rpt\_family="Alu"  
13546..14276  
/rpt\_family="Alu"  
complement(14646..14917)  
repeat\_region  
/rpt\_family="Alu"  
complement(15040..15200)  
/rpt\_family="Alu"  
15289..15571  
/rpt\_family="Alu"  
16315..16437  
/rpt\_family="Alu"  
complement(16471..16746)  
repeat\_region  
/rpt\_family="Alu"  
complement(16780..16914)  
/rpt\_family="MER21"  
16905..17072  
/product="Human ferritin heavy chain mRNA"  
17331..17503  
/rpt\_family="Alu"  
complement(17959..18124)  
repeat\_region  
/rpt\_family="MER21"  
18889..19131  
/rpt\_family="Alu"  
23180..23501  
/rpt\_family="Alu"  
complement(23733..24024)  
repeat\_region  
/rpt\_family="Alu"  
complement(24646..24717)  
/rpt\_family="MIR"  
24775..25060  
/rpt\_family="Alu"  
26119..26366  
/rpt\_family="Alu"  
complement(26555..26616)  
repeat\_region  
/rpt\_family="Alu"  
26767..27818  
/rpt\_family="Alu"  
complement(28535..28821)  
repeat\_region  
/rpt\_family="Alu"  
28895..29109  
/rpt\_family="MIR"  
29148..29439  
/rpt\_family="Alu"  
31140..31278  
/rpt\_family="MER4"  
31282..31675  
/rpt\_family="LTR7"  
31793..32411  
/rpt\_family="Alu"  
complement(32903..33202)  
repeat\_region  
/rpt\_family="Alu"  
complement(33214..33708)  
/rpt\_family="LTR8"  
complement(33709..33804)  
repeat\_region  
/rpt\_family="LTR9"  
33826..34102  
/rpt\_family="Alu"  
complement(34112..34172)  
repeat\_region  
/rpt\_family="LTR9"  
34470..34614  
/rpt\_family="Alu"

```
repeat_region 34913. .38086
/rpc_family="I1"
repeat_region 34994. .35249
/rpc_family="MER25"
repeat_region 36889. .37180
/rpc_family="Alu"
repeat_region 38416. .38688
/rpc_family="Alu"
repeat_region 38864. .39464
/rpc_family="Alu"
repeat_region 39952
complement(39826)
/rpc_family="Alu"
repeat_region 40060. .40131
complement(40060)
/rpc_family="MER4C"
repeat_region 40264. .40561
complement(40264)
/rpc_family="Alu"
repeat_region 40575. .40614
/rpc_family="Alu"
repeat_region 41866. .42139
/rpc_family="Mariner1"
repeat_region 42633. .42885
/rpc_family="Alu"
repeat_region 42948. .43175
/rpc_family="Alu"
repeat_region 43360. .43497
/rpc_family="Alu"
repeat_region 43744. .43924
/rpc_family="MLT1"
```

Query Match Best Local Similarity 70.5%; Score 1561.6; DB 2; Length 106648;  
Matches 1578; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 628 ACATGATCCTTATGAGGCTTGGAAACAAGTCTGTCTCTCACTGAACCAAGGGGCACT 687
DB 106648 ACATGATCCTTATGAGGCTTGGAAACAAGTCTGTCTCTCACTGAACCAAGGGGCACT 106589
QY 688 GCATTCCTGTGCCCCACCCAGCTTCCCTGTCTGTAGAGCTACCCCTTCTCCACAATT 747
DB 106588 GCATTCCTGTGCCCCACCCAGCTTCCCTGTCTGTAGAGCTACCCCTTCTCCACAATT 106529
QY 748 TCTTAGGGTTCATCATCGCGAGCAACATCGACCTAGGCCAGCACTGCTTGGGGTA 807
DB 106528 TCTTAGGGTTCATCATCGCGAGCAACATCGACCTAGGCCAGCACTGCTTGGGGTA 106469
QY 808 TATACCTGGCCACTTCAACAGGATCTTAGGGAAGTGTGGAGCCTTTCTCACTTAC 867
DB 106468 TATACCTGGCCACTTCAACAGGATCTTAGGGAAGTGTGGAGCCTTTCTCACTTAC 106409
QY 868 CTTGATCACCCGGGAAACTTTTGGGACCAAGTGAAGAGATGAGTTGTGTGAC 927
DB 106408 CTTGATCACCCGGGAAACTTTTGGGACCAAGTGAAGAGATGAGTTGTGTGAC 106349
QY 928 AGAATGCTGTGAGAACTGCCCCAGGGCTGACAGGCGAGGCTTACTGTAGCAATTTA 987
DB 106348 AGAATGCTGTGAGAACTGCCCCAGGGCTGACAGGCGAGGCTTACTGTAGCAATTTA 106289
QY 988 TCACTGGCCCCCACTTCTTTAGCAAGGGTGTGCTGACCAAAACCATGAGTGGCAATC 1047
DB 106288 TCACTGGCCCCCACTTCTTTAGCAAGGGTGTGCTGACCAAAACCATGAGTGGCAATC 106229
QY 1048 AGCTGATGACAGATGAACTTCCCACTAATTAATTTAGGGTATGAGTGAAGCACTTCA 1107
DB 106228 AGCTGATGACAGATGAACTTCCCACTAATTAATTTAGGGTATGAGTGAAGCACTTCA 106169
QY 1108 GGAAGAGGTGGAGGAAGTCTCTCACTCTAGAACTGTGATGATGATGAGTGAAGGTTG 1167
DB 106168 GGAAGAGGTGGAGGAAGTCTCTCACTCTAGAACTGTGATGATGATGAGTGAAGGTTG 106109
QY 1168 TCCAGGCGCTTATGATCTTTCTCACTGCTGGGAACTTCAACAAATACTTCTTCTT 1227
DB 106108 TCCAGGCGCTTATGATCTTTCTCACTGCTGGGAACTTCAACAAATACTTCTTCTT 106049
```

```
QY 1228 CTTGGGGGTCAGCCCAAGCTGTCACAAAATCAGATATTTCCCTTATTCAGATTTCT 1287
DB 106048 CTTGGGGGTCAGCCCAAGCTGTCACAAAATCAGATATTTCCCTTATTCAGATTTCT 105989
QY 1288 GACACTGTCAACCAATTTAACAACCCCACTTCAAGCCCAATCACTGTGGAGGAAGTGT 1347
DB 105988 GACACTGTCAACCAATTTAACAACCCCACTTCAAGCCCAATCACTGTGGAGGAAGTGT 105929
QY 1348 AACTTCCCTTTCTGAGATTTCTCAAGCATTTCTTCAAGGGTCAAGACGAGCTATTA 1407
DB 105928 AACTTCCCTTTCTGAGATTTCTCAAGCATTTCTTCAAGGGTCAAGACGAGCTATTA 105870
QY 1408 TGATGAAACCTTAAAGGCAACAATTCATTCCTGCTTCAAGCTAAGACAGACTT 1467
DB 105869 TGATGAAACCTTAAAGGCAACAATTCATTCCTGCTTCAAGCTAAGACAGACTT 105810
QY 1468 GGCAAAATCTGTGGCTCTTTCAGCAAGAGATGTTCAATTTAAGATCTTGTGGGC 1527
DB 105809 GGCAAAATCTGTGGCTCTTTCAGCAAGAGATGTTCAATTTAAGATCTTGTGGGC 105750
QY 1528 TGGGTGTGAGGCAAGTGAATCAAGAGGTGAGAGTTTGAACCACTGGCCAACT 1587
DB 105749 TGGGTGTGAGGCAAGTGAATCAAGAGGTGAGAGTTTGAACCACTGGCCAACT 105690
QY 1588 GATGAAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCGTGTGGTCTGT 1647
DB 105689 GATGAAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCGTGTGGTCTGT 105630
QY 1648 AGTCCCAAGCGAGAGGTGAGGGGAGATGCTTGAACCAAGAGTGTGGTGGCT 1707
DB 105629 AGTCCCAAGCGAGAGGTGAGGGGAGATGCTTGAACCAAGAGTGTGGTGGCT 105570
QY 1708 GAGATTTAGCACTGCAATCCAGCTGGGCGACGAGTGTGAGCTGTCTCAAAAAA 1767
DB 105569 GAGATTTAGCACTGCAATCCAGCTGGGCGACGAGTGTGAGCTGTCTCAAAAAA 105510
QY 1768 AAAAAGATCTCTCAACTTTGGCCCTCTACCTGCAACATTTGTGATTTGAATGA 1827
DB 105509 AAAAAGATCTCTCAACTTTGGCCCTCTACCTGCAACATTTGTGATTTGAATGA 105450
QY 1828 TACCTTCATCTTATGCTGTATATCTTTCATTCCTCACTAGGATGAGCACTTCTTA 1887
DB 105449 TACCTTCATCTTATGCTGTATATCTTTCATTCCTCACTAGGATGAGCACTTCTTA 105390
QY 1888 ACTGCTTCCATTAAGGGAGGCTTCCGACGAGGCTCTCAGCTGTGTGTACACGAG 1947
DB 105389 ACTGCTTCCATTAAGGGAGGCTTCCGACGAGGCTCTCAGCTGTGTGTACACGAG 105330
QY 1948 GACACTGATCAGTCAACGCAATACAGCTGTCCACCTGAGAGAGCTGTCTTAACAAC 2007
DB 105329 GACACTGATCAGTCAACGCAATACAGCTGTCCACCTGAGAGAGCTGTCTTAACAAC 105270
QY 2008 CTGAATCAATAGTTTATGCTTAATATATTAATAAATCCAGACTTACCTTAATGCT 2067
DB 105269 CTGAATCAATAGTTTATGCTTAATATATTAATAAATCCAGACTTACCTTAATGCT 105210
QY 2068 TTTATTCATTAATACTGTGAAAGCTAGACTGAACCATTTGGAATAATTTAAGTCT 2127
DB 105209 TTTATTCATTAATACTGTGAAAGCTAGACTGAACCATTTGGAATAATTTAAGTCT 105150
QY 2128 GATTTCAAGTGGGAGACCTTATGTTCTATCTGAATCCAAGACGACACTTAATTA 2187
DB 105149 GATTTCAAGTGGGAGACCTTATGTTCTATCTGAATCCAAGACGACACTTAATTA 105090
QY 2188 CTGCCAAACTAATGATTTAATAAATAC 2216
DB 105089 CTGCCAAACTAATGATTTAATAAATAC 105061
```

RESULT 11  
BC041664 2435 bp mRNA linear PRI 07-OCT-2003  
LOCUS BC041664  
DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease),

beactrophin), mRNA (CDNA clone MGC:47884 IMAGE:5194649), complete cda.

ACCESSION BC041664  
VERSION BC041664.1 GI:27371319  
KEYWORDS MGC.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 2435)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Parker, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., Moewan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.R., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kesteman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywicki, M.I., Skalska, U., Smalins, D.E., Schnerch, A., Schein, J.E., Jones, S.J., Skalska, U., Smalins, D.E., human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE 2 (bases 1 to 2435)  
Strausberg, R.  
Direct Submission  
Submitted (20-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

JOURNAL NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
MEDLINE Contact: MGC help desk  
PUBMED Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

REMARK COMMENT  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland.  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nigrl.nih.gov](mailto:nisc_mgc@nigrl.nih.gov)  
Aahter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R., Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McElwee, J.C., McDowell, J., Pearson, R., Stancirpop, S., Thomas, P.J., Touchman, J.W., Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

FEATURES  
source  
1. .2435  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:47884 IMAGE:5194649"  
/issue\_type="Brain, adult, 6 pooled whole brains"  
/clone\_id="NIH MGC\_114"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"

gene 1. .2435  
/gene="VMD2"  
/note="synonyms: BMD, BEST"  
/db\_xref="locusid:7439"  
/db\_xref="MIM:153700"  
87. .1901  
/codon\_start=1  
/product="VMD2 protein"  
/protein\_id="AAH41664.1"  
/db\_xref="GI:27371320"  
/translation="MEKTLVDSYIOLIPISFVLGFVTLVVTTRMMQYENLPMPDRLMSLVSGVEGDEOGRLLRLIRYANGLVILRSVSTAYKRPSSQNLVQAFMPRAEKKOLKSLSPNNMFWPVMVPMVSMKMLGRIIDPILLOSILNEMTLRPOCGHLYADNISIPLYQVTVAVYFPLTCLVGRQPLAKAYRGNEDLYVVPVPLQFETVGMKVAESQILNPFGEEDDFETVIRLQVSLVADEMDQILPMBEDMYNMRPPOPEPTAASQFRRVSGSTFNI SLNKEBMEQPNODEBDAIIGRLVGLQSHNDPRRANRRTKLPKPKESLHLBGLPQNHKAQNVGDEBDAIIGRLVADFKAPLYQRPQVYASAPQPLSPFPMFPELPSPSKLSVTGIDTQKSLKTVSSGAKKSPFLSESDGALMEHREVSQVKTVEFNLTPMPTIPENHLKEPLQSPNTHTTLKDHNDPYWALNRSVTLNNOCHTALCPRTASLALSLPLNLFGRHNCOSTLDRRLAMGIVATFTGLIGKCSGRLTSPVTHREDFLGRGR"

misc\_feature  
/note="DUF289; Region: Putative membrane protein. This family called family 8 in. may be a transmembrane protein The specific function of this protein is unknown"  
/db\_xref="CD:pfam01062"

ORIGIN  
Query Match 60.5%; Score 1339.8; DB 9; Length 2435;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1341; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
17 CAGCTGGAACAAAGAGATGAGTTCAGGCCAATCAGAGACGAGAGATGCTCA 76  
1004 CAGCTGGAACAAAGAGATGAGTTCAGGCCAATCAGAGAGACGAGAGATGCTCA 1063  
77 CGCTGCATCATTTGGCGCTCTTCTAGCCGCGAGTCCAGATACCATCTCCAGGGC 136  
1064 CGCTGCATCATTTGGCGCTCTTCTAGCCGCGAGTCCAGATACCATCTCCAGGGC 1123  
137 AAACCTAAGGACCAAACTACTGTGGCCCAAGAGGAACTCCCTCTCCAGAGGGCTGCC 196  
1124 AAACCTAAGGACCAAACTACTGTGGCCCAAGAGGAACTCCCTCTCCAGAGGGCTGCC 1183  
197 CAAAAACCAAGGACCAAACTACTGTGGCCCAAGAGGAACTCCCTCTCCAGAGGGCTGCC 256  
1184 CAAAAACCAAGGACCAAACTACTGTGGCCCAAGAGGAACTCCCTCTCCAGAGGGCTGCC 1243  
257 GCTTAAAGGCTGTGAGGCTCTTCAAGTCTGCCCACTGTATCAAGGCTAGCTACTAG 316  
1244 GCTTAAAGGCTGTGAGGCTCTTCAAGTCTGCCCACTGTATCAAGGCTAGCTACTAG 1303  
317 TGCCCAAGAGAGCCCTCAGCCCACTCCCATGTTCTTCCCTAGAACCATCAGGGCC 376  
1304 TGCCCAAGAGAGCCCTCAGCCCACTCCCATGTTCTTCCCTAGAACCATCAGGGCC 1363  
377 GTCAAAGCTTACAGGTGTACAGGCAATGACCAACCAAGCAAAAGCTTAAAGCTGTAG 436  
1364 GTCAAAGCTTACAGGTGTACAGGCAATGACCAACCAAGCAAAAGCTTAAAGCTGTAG 1423  
437 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGAGAGGGGCTTGTAGTGA 496  
1424 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGAGAGGGGCTTGTAGTGA 1483  
497 CCAGAGATATCTCAAGTGAAGAGAGAACTGTGAGATTAACTGACGGATATGCGAGA 556  
1484 CCAGAGATATCTCAAGTGAAGAGAGAACTGTGAGATTAACTGACGGATATGCGAGA 1543  
557 GATCCCGGAAATCAGCTCAAGAACCTTTGGAACATCAACCAACATACACTAC 616  
1544 GATCCCGGAAATCAGCTCAAGAACCTTTGGAACATCAACCAACATACACTAC 1603

[illegible]

```

    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11q12-q13.1"
    /clone="PAC dJ466a11"
    36..674
    /gene="VMD2"
    /number=10

```

17 ATCTCTCGTTCTTCCAGCTGACCTGAACAAGAGAGATGAGTTCCAGCCCAATCAGAGG 60

121 ACCATCTCCCGGGCAAACTCAAGAGCAAACTACTGTGCTCCAGAGGCTTCTTCTCTTCT 136

197 TCCAGAGGGCTGCCAAAACCAAGGCGCCAAACAGACGTTAGGGGCCAGGAAG 240

25/ ACACACAGGCGCTGAAAGCTTAAAGGCTGTGACCGCTTCAAGTCTTGCCCACTGTATTACAGA 316

Y  
361 TAGAACCATCAGCGCCGTCAAAGCTTCACAGTGCACAGCATATACACCATAAGACAAA 420  
377 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT

437 GCTTAAAGACTGAGTTCTGGGGCCAGAAAGTTTGATTGCTCTCAGAGAGCGATG 480

437 GGGCTTGAAGGACCCAGAAAGTATCTCAAGTAGGAGGAGAAACTGTGAGATTAAAC 556

557 TGACGGATATGCCAGAGATCCCCCGAATTTCGCTTCC  
600

QY 601 CCAACATACACTACACTCACTCAAGATGATGATCTTATTGGGCTTGGAAAAACAGT 660  
 DB 617 CCAACATACACTACACTCAAGATGATGATCTTATTGGGCTTGGAAAAACAGT 676  
 QY 661 CTGTCTCTCACTGACCAAGGAGGAGCTGCA 690  
 DB 677 CTGTCTCTCACTGACCAAGGAGGAGCTGCA 706

RESULT 13  
 AY357926S10 726 bp DNA linear PRI 26-AUG-2003  
 LOCUS AY357926S10  
 DEFINITION Macaca fascicularis beestrophin (VMD2) gene, exon 10.  
 ACCESSION AY357935  
 VERSION AY357935.1 GI:34013779  
 KEYWORDS 10 of 11  
 SEGMENT Macaca fascicularis (crab-eating macaque)  
 SOURCE Macaca fascicularis  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.  
 REFERENCE 1 (bases 1 to 726)  
 AUTHORS Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and Iwata, T.  
 TITLE Molecular cloning of VMD2 gene from Cynomolgus monkey (Macaca fascicularis)  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 726)  
 AUTHORS Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and Iwata, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-AUG-2003) National Institute of Sensory Organs, National Tokyo Medical Center, 2-5-1, Higashi-agaoka, Meguro-Ku, Tokyo 152-0021, Japan  
 FEATURES  
 source  
 1..726  
 /organism="Macaca fascicularis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9541"  
 36..674  
 /gene="VMD2"  
 /number=10

exon

ORIGIN

Query Match 29.9%; Score 662; DB 9; Length 726;  
 Best Local Similarity 95.8%; Pred. No. 2.1e-193;  
 Matches 680; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 1 ATCTCCGTTCTTTCCAGGCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGG 60  
 DB 17 ATCTCCGTTCTTTCCAGGCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGG 76  
 QY 61 ACAGAGAGATGCTCAAGCTGATCATTTGGCCGCTTCTAGGCGCTGCACTCCATGATC 120  
 DB 77 ACAGAGAGATGCTCAAGCTGATCATTTGGCCGCTTCTAGGCGCTGCACTCCATGATC 136  
 QY 121 ACCATCTCTCCAGGAGCAATCAAGACCAACTACTGTGGCCCAAGAGGAATCCCTTC 180  
 DB 137 ACCATCTCTCCAGGAGCAATCAAGACCAACTACTGTGGCCCAAGAGGAATCCCTTC 196  
 QY 181 TCCAGAGGAGCTGAGCCCAAAAACCAAGAGGAGCAAGAAAGGTTAGGGGCGAGGAAG 240  
 DB 197 TCCAGAGGAGCTGAGCCCAAAAACCAAGAGGAGCAAGAAAGGTTAGGGGCGAGGAAG 256  
 QY 241 ACAACAGAGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 300  
 DB 257 ACAACAGAGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 316  
 QY 301 GGCAGAGCTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360  
 DB 317 GGCAGAGCTTACATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 376

QY 361 TAGAACATCAGGAGCGGCTCAAAAGCTTCAACAGTGTCAAGGATAGACACCAAGACAAA 420  
 DB 377 CAGAACATCAGGAGCGGCTCAAAAGCTTCAACAGTGTCAAGGATAGACACCAAGACAAA 436  
 QY 421 GCTTAAAGACTGTGAGTCTGTGGGAGCAAGAAAGTTTGAATGCTCTCAGAGAGGATG 480  
 DB 437 GCTTAAAGACTGTGAGTCTGTGGGAGCAAGAAAGTTTGAATGCTCTCAGAGAGGATG 496  
 QY 481 GGGGCTTGTGAGGAGCAAGGAGTATCTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 DB 497 GGGGCTTGTGAGGAGCAAGGAGTATCTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGG 556  
 QY 541 TGAAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 DB 557 TGAAGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 616  
 QY 601 CCAACATACACTACACTCAAGATGATGATGATCTTATTGGGCTTGGAAAAACAGT 660  
 DB 617 CCAACATACACTACACTCAAGATGATGATGATCTTATTGGGCTTGGAAAAACAGT 676  
 QY 661 CTGTCTCTCACTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720  
 DB 677 CTGTCTCTCACTGACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 726

RESULT 14  
 BD132659 1263 bp DNA linear PAT 18-SEP-2002  
 LOCUS BD132659  
 DEFINITION Secreted proteins and polynucleotides encoding chem.  
 ACCESSION BD132659  
 VERSION BD132659.1 GI:23227604  
 KEYWORDS JP 2002504822-A/2  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1263)  
 AUTHORS Jacobs, R., McCoy, J.M., Lavallie, E.R., Racie, L.A., Treac, M., Spaulding, V., Agostino, M.J., Hones, S.H. and Fechtel, K.  
 TITLE Secreted proteins and polynucleotides encoding them  
 JOURNAL Patent: JP 2002504822-A 2 12-FEB-2002;  
 COMMENT GENETICS INSTITUTE INC  
 PN JP 2002504822-A/2  
 PD 12-FEB-2002  
 PR 08-JUN-1998 JP 199503038  
 PR 11-JUN-1997 US 08/873218, 05-JUN-1998 US 09/092722 PI  
 KENNETH JACOBS, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI MAURICE TREACY,  
 PI VIRKI SPAULDING, MICHAEL J AGOSTINO, STEVEN H HONES, KIM FECHTEL,  
 PC C12N15/12, C07K14/47, A61K38/17  
 CC Strandedness: Double;  
 CC Topology: Linear;  
 FH Key  
 Location/Qualifiers  
 source  
 1..1263  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

ORIGIN

Query Match 29.1%; Score 643.8; DB 6; Length 1263;  
 Best Local Similarity 99.7%; Pred. No. 1e-187;  
 Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 CAGCTGAACCAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGAGATGCTCA 76  
 DB 225 CAGCTGAACCAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGAGATGCTCA 284  
 QY 77 CGTGGATCATTTGGGCGGCTTCTTGAAGCTGTGAGTCCCATGATCAACATCTTCCAGGGC 136  
 DB 285 CGTGGATCATTTGGGCGGCTTCTTGAAGCTGTGAGTCCCATGATCAACATCTTCCAGGGC 344  
 QY 137 AAACCTCAAGAGACCAACTACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 196

```

Db      345 AAACCTCAGAGCAAACTACTGTGTGCCCCAAGAGGGAATCCCTTCTCCACAGAGGCGCTGCCC
QY      197 CAAAACCAACAGAGCAGCCCAACAGAACTTAAAGGCGCAGAGAGACAAAGGCTTGAA
Db      405 CAAAACCAACAGAGCAGCCCAACAGAACTTAAAGGCGCAGAGAGACAAAGGCTTGAA
QY      257 GCTTAAAGCTGTGAGAGCCTTCAAGTGTGCCCCAAGTGTATCAAGGCGAGGCTACTAG
Db      465 GCTTAAAGCTGTGAGAGCCTTCAAGTGTGCCCCAAGTGTATCAAGGCGAGGCTACTAG
QY      317 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAAACATTCAGCCGC
Db      525 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAAACATTCAGCCGC
QY      377 GTCAAGGCTTCAAGTGTCAAGGCAATACACCAAGCAAAAGCTTAAAGACTGTGAG
Db      585 GTCAAGGCTTCAAGTGTCAAGGCAATACACCAAGCAAAAGCTTAAAGACTGTGAG
QY      437 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGCA
Db      645 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGGCTTGATGAGCA
QY      497 CCCAAGAGTATCTCAAGTGTGAGAGAACTGTGAGTTAACTTACGAGATATGCCAGA
Db      705 CCCAAGAGTATCTCAAGTGTGAGAGAACTGTGAGTTAACTTACGAGATATGCCAGA
QY      557 GATCCCCGAAATTCACCTCAAGAAACCTTGGAAACATCCAAACCAATACACATCAC
Db      765 GATCCCCGAAATTCACCTCAAGAAACCTTGGAAACATCCAAACCAATACACATCAC
QY      617 ACTCAAAAGTCAATGATCTTATTGAGGCTTGGAAAAAGTGTG 663
Db      825 ACTCAAAAGTCAATGATCTTATTGAGGCTTGGAAAAAGGAGTGT 871

RESULT 15
AF057169      2210 bp      mRNA      linear      PRI 17-OCT-1998
LOCUS      Homo sapiens beestrophin (VMD2) mRNA, alternatively spliced product,
ACCESSION      AF057169
VERSION      AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Petrushkin, K., Koisti, M.J., Bakall, B., Li, W., Xie, G., Marknell, T.,
            Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
            Metzger, M.L., Caskey, C.T. and Wadelius, C.
            Identification of the gene responsible for Best macular dystrophy
            Nat. Genet. 19 (3), 241-247 (1998)
TITLE
JOURNAL      96324772
MEDLINE      96623395
PUBMED      2 (bases 1 to 2210)
            Petrushkin, K.
AUTHORS      Direct Submission
            Submitted (03-APR-1998) Human Genetics, Merck Research
            Laboratories, West Point, PA 19486, USA
FEATURES
SOURCE
            1. .2210
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="11"
            1. .2210
            /gene="VMD2"
            105. .1862
            /gene="VMD2"

```

```

/note="Best's macular dystrophy gene; BMD; alternatively
splice isoform"
/codon_start=1
/product="beestrophin"
/protein_id="AAC64343.1"
/db_xref="GI:3335159"
/translation="MTITVTSOVANRILGFSRLLICMRGSIYKLYKGFLLFLCY
IIRPIYALTEEOILMEPKTLTYCDSYIQLPISFVLSFVTLVTRWNOYENLPW
PDRILMSVGFVSKDEOGRILRLIRLANLGNVILSVSTAYKRPSPAOHLVOA
GHPTRPBNKOLEKLSLPHNFWPVMWNLNLMKMLGRIIDPILLOSLANENMLR
TQGGHLYADWISIPLYVQVYVAVSFFLCVGRGCPRIKADYVGEHDLVYVVF
TLQFFVYGMVLKVAEQLINPGEEDDDDEFTWMIYDRNLQVSLAVDEHODLPMPER
DMYMNKDEPQRYTAASQFRASFPWGSFTNLSIKKEWEPQPNODEBDAAGITGR
FLGLSHDHPHPRANSRTLLMPKPSLHBEGLPKNHKAKQNVGQEDNKAUKLAV
DAKSAPLYORBGVYSAPOTPSLPTFMFPPLSPASKUSVYGLDIDKSLKTVSSG
AKSFPLBSDSGLMHPVSVQVRKTFVFNLTMPPEIPENHMLEPLEOSPNTIHTT
LNQHMDPYVALENRDEAHS"

```

## ORIGIN

```

Query Match      29.1%; Score 643.8; DB 9; Length 2210;
Best Local Similarity 99.7%; Pred. No. 1.1e-187;
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      17 CAGCTGAAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGAGATGCTCA 76
Db      1202 CAGCTGAAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
QY      77 CGCTGCATCATTTGGCGCGCTCTCTAGGCGCTGAGCCCATGATCACTCTCCAGAGGC
Db      1262 CGCTGCATCATTTGGCGCGCTCTCTAGGCGCTGAGCCCATGATCACTCTCCAGAGGC
QY      137 AAACCTCAAGACCAACTGCTGTGCCCCAAGAGGAATCCCTTCCACGAGGCGCTGCC
Db      1322 AAACCTCAAGACCAACTGCTGTGCCCCAAGAGGAATCCCTTCCACGAGGCGCTGCC
QY      197 CAAAACCAACAGAGCAGCCCAACAGAACTTAAAGGCGCAGAGAAACAAACAGGCTTGAA
Db      1382 CAAAACCAACAGAGCAGCCCAACAGAACTTAAAGGCGCAGAGAAACAAACAGGCTTGAA
QY      257 GCTTAAAGCTGTGAGAGCGCTTCAAGTGTGCCCCCACTGATCAGAGGCGAGCTTCAAG
Db      1442 GCTTAAAGCTGTGAGAGCGCTTCAAGTGTGCCCCCACTGATCAGAGGCGAGCTTCAAG
QY      317 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAAACATTCAGCGCC
Db      1502 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTAAACATTCAGCGCC
QY      377 GTCAAAAGCTTCAAGTGTCAAGGCAATACACCAAGCAAAAGCTTAAAGCTGTGAG
Db      1562 GTCAAAAGCTTCAAGTGTCAAGGCAATACACCAAGCAAAAGCTTAAAGCTGTGAG
QY      437 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGTAGAGCA
Db      1622 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCGCTTGTAGAGCA
QY      497 CCCAAGAGTATCTCAAGTGTGAGAGAAACTGTGAGTTAACTTACCGAGATATGCCAGA
Db      1682 CCCAAGAGTATCTCAAGTGTGAGAGAAACTGTGAGTTAACTTACCGAGATATGCCAGA
QY      557 GATCCCCGAAATTCACCTCAAGAAACCTTGGAAACATCCAAACCAATACACATCACATCAC
Db      1742 GATCCCCGAAATTCACCTCAAGAAACCTTGGAAACATCCAAACCAATACACATCACATCAC
QY      617 ACTCAAAAGTCAATGATCTTATTGAGGCTTGGAAAAAGTGTG 663
Db      1802 ACTCAAAAGTCAATGATCTTATTGAGGCTTGGAAAAAGGAGTGT 1848

```

Search completed: March 25, 2004, 05:47:07  
Job time : 8608 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 01:49:44 ; Search time 855 Seconds

(without alignments)  
11010.534 Million cell updates/sec

Title: US-09-622-964-1\_COPY\_13900\_16115

Perfect score: 2216  
Sequence: 1 actctctgtctcttcacgc.....ctaatgagtttaataataac 2216

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N\_Geneseq\_29Jan04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2003as: \*  
8: geneseqn2003bs: \*  
9: geneseqn2003cs: \*  
10: geneseqn2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2215	100.0	16125	2	AAZ21226	Aaz21226 Human CGI
2	2188.6	98.8	16650	5	ABAI4559	Abai4559 Human ner
3	2188.6	98.8	18530	5	ABAI4557	Abai4557 Human ner
4	2187	98.7	18537	5	ABAI4558	Abai4558 Human ner
5	1631.6	73.6	1717	6	ABLB9697	AbLB9697 Human pol
6	643.8	29.1	1253	2	AAV99722	Aav99722 Human adu
7	642.2	29.0	1758	7	ABZ80972	AbZ80972 Human bee
8	642.2	29.0	2229	2	AAZ21227	Aaz21227 Human CGI
9	642.2	29.0	2429	2	AAZ21228	Aaz21228 Human CGI
10	511.8	22.3	1238	8	ADAA4960	Ada44960 Human pol
11	494	22.3	7108	8	ABAI4556	Abai4556 Human ner
12	404.4	18.2	459	8	ACH39717	Ach39717 Human foe
13	399.4	18.0	424	2	AAV86976	Avv86976 EST clone
14	339.4	15.3	1326	7	ACA03889	Aca03889 cDNA upre
15	318	14.4	356	3	AAC20667	Aac20667 Human sec
16	164	7.4	1119	7	ABZ71972	Abz71972 Human fer
17	164	7.4	1198	6	ABK83853	Abk83853 Human cdn
18	164	7.4	1198	6	ABN95676	Abn95676 Gene #217
19	164	7.4	1234	5	AA667682	Aa667682 DNA encod
20	164	7.4	1279	5	AA685665	Aa685665 DNA encod
21	164	7.4	2183	5	AA691587	Aa691587 DNA encod
22	133.2	6.0	65042	7	AB855711	Ab855711 DNA encod
23	133.2	6.0	65042	8	ACA63029	Aca63029 Human gen

24	133.2	6.0	149480	6	ABL61947	AbL61947
25	133.2	6.0	149480	6	ABL68365	AbL68365
26	133.2	6.0	149480	6	ABL61948	AbL61948
27	127.6	5.8	322101	9	AAD58431	Aad58431
28	126.8	5.7	22900	4	AAK82210	Aak82210
29	126.2	5.7	22645	4	AAK89727	Aak89727
30	126.2	5.7	22645	4	AAI05495	Aai05495
31	126.2	5.7	22645	4	AAI04985	Aai04985
32	126.2	5.7	22645	4	ABL97878	AbL97878
33	126.2	5.7	22645	4	ABL98348	AbL98348
34	124.8	5.6	165199	6	ABK83460	AbK83460
35	124.8	5.6	177531	7	ACF62732	AcF62732
36	124.8	5.6	177531	7	ADB20847	AdB20847
37	124.8	5.6	177531	9	ADB87936	AdB87936
38	124.8	5.6	177531	9	ADB87936	AdB87936
39	124.8	5.6	177531	9	ADB86919	AdB86919
40	124.2	5.6	38997	6	AAH36069	Aah36069
41	124	5.6	700	4	AAH92912	Aah92912
42	124	5.6	96587	8	ADA02984	Ada02984
43	124	5.6	96587	9	ADB72722	AdB72722
44	124	5.6	96587	9	ADC85464	AdC85464
45	123.8	5.6	14874	4	AAK89726	Aak89726

## ALIGNMENTS

RESULT 1	
AAZ21226	AAZ21226 standard; DNA; 16125 bp.
XX	
AC	AAZ21226;
XX	
DT	22-NOV-1999 (first entry)
XX	
DE	Human CGI
XX	
KW	CGI, Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW	age-related macular dystrophy; ss.
XX	
XX	Homo sapiens.
OS	
XX	
FH	Key
FT	intron
FT	Location/Qualifiers
FT	1. 1513
FT	/*tag= a
FT	/number= 1
FT	1514. 1570
FT	/*tag= b
FT	/number= 1
FT	1571. 3278
FT	/*tag= c
FT	/number= 2
FT	3279. 3466
FT	/*tag= d
FT	/number= 2
FT	3315. 15787
FT	/*tag= e
FT	/product= "CGI
FT	/note= "contains introns"
FT	3467. 6616
FT	/*tag= f
FT	/number= 3
FT	6617. 6711
FT	/*tag= g
FT	/number= 3
FT	6712. 7227
FT	/*tag= h
FT	/number= 4
FT	7228. 7461
FT	/*tag= i
FT	/number= 4
FT	7462. 8353
FT	/*tag= j
FT	intron



FT /number= 5  
FT exon 8354..8508  
FT /tag= k  
FT /number= 5  
FT intron 8509..8896  
FT /tag= 1  
FT /number= 6  
FT exon 8897..8974  
FT /tag= m  
FT /number= 6  
FT intron 8975..9676  
FT /tag= n  
FT /number= 7  
FT exon 9677..10032  
FT /tag= o  
FT /number= 7  
FT intron 10033..11212  
FT /tag= p  
FT /number= 8  
FT exon 11213..11293  
FT /tag= q  
FT /number= 8  
FT intron 11294..11556  
FT /tag= r  
FT /number= 9  
FT exon 11557..11708  
FT /tag= s  
FT /number= 9  
FT intron 11709..13918  
FT /tag= t  
FT /number= 10  
FT exon 13919..14557  
FT /tag= u  
FT /number= 10  
FT intron 14558..15768  
FT /tag= v  
FT /number= 11  
FT exon 15769..16125  
FT /tag= w  
FT /number= 11  
PN WO9943695-A1.

XX 02-SEP-1999.  
XX 22-FEB-1999; 99WO-US003790.  
XX 25-FEB-1998; 98US-0075941P.  
XX 18-DEC-1998; 98US-0112926P.  
XX (MERI) MERCK & CO INC.  
XX (UYUP-) UNIV UPSALA.  
XX Petrukhin K, Caskey CT, Metzker M, Wadelius C;  
XX WPI; 1999-540560/45.  
XX P-PSDB; AA129953.  
XX Human and mouse polynucleotides encoding CGICE polypeptides.  
XX Claim 2; Fig 1; 67pp; English.

XX The present sequence represents the human CGICE gene, which when mutated  
XX is responsible for Best's macular dystrophy (BMD). Polynucleotides  
XX encoding CGICE are useful for diagnosing whether a patient carries a  
XX mutation in the CGICE gene. Normal and mutated CGICE proteins are useful  
XX for identifying activators and/or inhibitors of these proteins, in order  
XX to treat BMD. The CGICE gene offers a simpler and cheaper method of  
XX diagnosing BMD without the need for the presence of the patient. The gene  
XX may also be useful to discovering the genetic cause of age-related  
XX macular dystrophy

XX Sequence 16125 BP; 3988 A; 4175 C; 4215 G; 3726 T; 0 U; 21 Other;

Query Match 100.0%; Score 2215; DB 2; Length 16125;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCTCTGTTCTTCTTCCAGCTGAACTGAGAGAGATGAGATTCAGCCCAATCAGAGAG 60  
Db 13900 ATCTCTGTTCTTCTTCCAGCTGAACTGAGAGAGATGAGATTCAGCCCAATCAGAGAG 13959  
QY 61 ACAGAGAGATGCTCAGCTGCGATCATTTGGCGGCTTCTTGGAGCTTCCATGATC 120  
Db 13960 ACAGAGAGATGCTCAGCTGCGATCATTTGGCGGCTTCTTGGAGCTTCCATGATC 14019  
QY 121 ACCATCTCTCCAGGCGAACTCAAGAGCAAACTACTGTGGCCCAAGAGAGATCCCTTC 180  
Db 14020 ACCATCTCTCCAGGCGAACTCAAGAGCAAACTACTGTGGCCCAAGAGAGATCCCTTC 14079  
QY 181 TCACAGAGGCTGCGCCCAAAAACACAGAGGCGCAACAGAGCTTGGGGCCAGAGAG 240  
Db 14080 TCACAGAGGCTGCGCCCAAAAACACAGAGGCGCAACAGAGCTTGGGGCCAGAGAG 14139  
QY 241 ACAACAGAGCTGGAAGCTTAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 300  
Db 14140 ACAACAGAGCTGGAAGCTTAAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAG 14199  
QY 301 GGCAGAGCTTACTACAGTGGCCCAAGAGGCGCCCTTCAAGCCCACTTCCATGTTCTTCCCC 360  
Db 14200 GGCAGAGCTTACTACAGTGGCCCAAGAGGCGCCCTTCAAGCCCACTTCCATGTTCTTCCCC 14259  
QY 361 TAGAACATCAGAGGCGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 420  
Db 14260 TAGAACATCAGAGGCGCTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAG 14319  
QY 421 GCTTAAAGCTGAGTCTTGGGGCCCAAGAGAAATTTGAATGCTTCTCAGAGAGCGATG 480  
Db 14320 GCTTAAAGCTGAGTCTTGGGGCCCAAGAGAAATTTGAATGCTTCTCAGAGAGCGATG 14379  
QY 481 GGGCTTGAATGAGACCCCAAGAAATATCTCAAGTGAAGAGAGAAATCTGAGAGTTTAACC 540  
Db 14380 GGGCTTGAATGAGACCCCAAGAAATATCTCAAGTGAAGAGAGAAATCTGAGAGTTTAACC 14439  
QY 541 TGACGAGATATGCGAGATATCCCGAAATATCACTCAAGAACTTTGAAATATCAACAA 600  
Db 14440 TGACGAGATATGCGAGATATCCCGAAATATCACTCAAGAACTTTGAAATATCAACAA 14499  
QY 601 CCAACATACACACTACACTCACTCAAGATCACTGATCTTATTTGGGCTTGGAGAAACAGGT 660  
Db 14500 CCAACATACACACTACACTCACTCAAGATCACTGATCTTATTTGGGCTTGGAGAAACAGGT 14559  
QY 661 CTGTCTCCAGCTGGAACCAAGGGGCACTGATGCTTGGGCGCCCAAGCCAGCTTCCCTTG 720  
Db 14560 CTGTCTCCAGCTGGAACCAAGGGGCACTGATGCTTGGGCGCCCAAGCCAGCTTCCCTTG 14619  
QY 721 CTCTGAGCTTACCTTCTCTCAAAATTTCTAGGGTTCATGATGCGAGAGCACTGCG 780  
Db 14620 CTCTGAGCTTACCTTCTCTCAAAATTTCTAGGGTTCATGATGCGAGAGCACTGCG 14679  
QY 781 ACCTAGGCCAGAGCTGGGCTTGGGTATATCTTGGGCCACTTCAAGAGATCTTAGAGA 840  
Db 14680 ACCTAGGCCAGAGCTGGGCTTGGGTATATCTTGGGCCACTTCAAGAGATCTTAGAGA 14739  
QY 841 AGTGTTCGAGACCTTTCTTCACTTCACTGATATCACTGAGAGACTTTCTTGGAGACAG 900  
Db 14740 AGTGTTCGAGACCTTTCTTCACTTCACTGATATCACTGAGAGACTTTCTTGGAGACAG 14799  
QY 901 GGAAGAGAGATGAGTGTGTCTGACAGAAATGCTGTGAGAGAACTGCCAGAGGCTGAC 960  
Db 14800 GGAAGAGAGATGAGTGTGTCTGACAGAAATGCTGTGAGAGAACTGCCAGAGGCTGAC 14859  
QY 961 AGGCGAGCTTAGCTGAGAGAGATGATATCACTGAGGCGCCCACTTACTTGAAGCAAGGCTG 1020  
Db 14860 AGGCGAGCTTAGCTGAGAGAGATGATATCACTGAGGCGCCCACTTACTTGAAGCAAGGCTG 14919



QY 1021 CTGACCCAAAACATGAGTGGCAGTGCAGTGTGATGACATGTAACACTTCCCCCAATAC 1080  
| | | | |  
Db 14920 CTGACCCAAAACATGAGTGGCAGTGCAGTGTGATGACATGTAACACTTCCCCCAATAC 14979  
| | | | |  
QY 1081 TATTTAGGGTAGTACCCAGACATACAGAAAAGGTGGCAGAAAGCTGCTCCTACTCTTAGG 1140  
| | | | |  
Db 14980 TATTTAGGGTAGTACCCAGACATACAGAAAAGGTGGCAGAAAGCTGCTCCTACTCTTAGG 15039  
| | | | |  
QY 1141 AACTGAGTGAATGGAGAGTGTGAGGGGTGTCCAGCGCCCTTGTAGTCATTTTCTCACTGCTG 1200  
| | | | |  
Db 15040 AACTGAGTGAATGGAGAGTGTGAGGGGTGTCCAGCGCCCTTGTAGTCATTTTCTCACTGCTG 15099  
| | | | |  
QY 1201 GGAACCTCACCCAAAATACTTCTTGCTTCTTGGGGGTCAAGCCAAAGCTGTACAAAATCA 1260  
| | | | |  
Db 15100 GGAACCTCACCCAAAATACTTCTTGCTTCTTGGGGGTCAAGCCAAAGCTGTACAAAATCA 15159  
| | | | |  
QY 1261 GATATTTTCCCTTTATTTCCAGATTTCTGACACTGTGACCCCAATTATTAACACCCCACTT 1320  
| | | | |  
Db 15160 GATATTTTCCCTTTATTTCCAGATTTCTGACACTGTGACCCCAATTATTAACACCCCACTT 15219  
| | | | |  
QY 1321 CAGCCCAATCAGTGGGAGAAAGTGTACTTCCCTTTTGTGATTTCTCAAGCAGTTACT 1380  
| | | | |  
Db 15220 CAGCCCAATCAGTGGGAGAAAGTGTACTTCCCTTTTGTGATTTCTCAAGCAGTTACT 15279  
| | | | |  
QY 1381 TTCACGGGTGAGAAACGACGCTATTATGATTGAAAACCTTAAAGGSCAACAAATTTCAAT 1440  
| | | | |  
Db 15280 TTCACGGGTGAGAAACGACGCTATTATGATTGAAAACCTTAAAGGSCAACAAATTTCAAT 15339  
| | | | |  
QY 1441 CTTGCTTTCTAGGCTTAAGACAGAACTTGGCAAACTGTGCTGTTCAGCAAAAGGATG 1500  
| | | | |  
Db 15340 CTTGCTTTCTAGGCTTAAGACAGAACTTGGCAAACTGTGCTGTTCAGCAAAAGGATG 15399  
| | | | |  
QY 1501 TTGATATTTAAAGATCTGTCTTGGGCTGGGTGTGGAGGCAAGTAATCAGAGAGTCA 1560  
| | | | |  
Db 15400 TTGATATTTAAAGATCTGTCTTGGGCTGGGTGTGGAGGCAAGTAATCAGAGAGTCA 15459  
| | | | |  
QY 1561 GAGATTGAGACCAACCTGGSCAAATGATGAAACCCCATCTCTCAACCAAAAATAATCA 1520  
| | | | |  
Db 15460 GAGATTGAGACCAACCTGGSCAAATGATGAAACCCCATCTCTCAACCAAAAATAATCA 15519  
| | | | |  
QY 1621 ATCAGCTGGCCGTGTGGTGTGCTGTAGTCCCAACGACAGAGGTTGAGGGGAGAAATGTC 1680  
| | | | |  
Db 15520 ATCAGCTGGCCGTGTGGTGTGCTGTAGTCCCAACGACAGAGGTTGAGGGGAGAAATGTC 15579  
| | | | |  
QY 1681 TTGAAACCCAGAGAGTGTGTGTCAGTGAATGAGCAATCTGCAATCTCAAGCTGGGGGAC 1740  
| | | | |  
Db 15580 TTGAAACCCAGAGAGTGTGTGTCAGTGAATGAGCAATCTGCAATCTCAAGCTGGGGGAC 15639  
| | | | |  
QY 1741 GAGTGTGAGCTGTCTCAAAAATAAAAGGATGCTCAACCTTGGCCCTCTACT 1800  
| | | | |  
Db 15640 GAGTGTGAGCTGTCTCAAAAATAAAAGGATGCTCAACCTTGGCCCTCTACT 15699  
| | | | |  
QY 1801 GCAACATTTTGTATTTGAATGAAGGTACCTTCCATACTTATGCTGTATTAATCTTTCAT 1860  
| | | | |  
Db 15700 GCAACATTTTGTATTTGAATGAAGGTACCTTCCATACTTATGCTGTATTAATCTTTCAT 15759  
| | | | |  
QY 1861 TCTCAGTGAAGGAGGACATCTTCACTGCTTCTTAATGGGATGCTTCCGACGCCA 1920  
| | | | |  
Db 15760 TCTCAGTGAAGGAGGACATCTTCACTGCTTCTTAATGGGATGCTTCCGACGCCA 15819  
| | | | |  
QY 1921 GGTCTCAGCTGTGTGTGACACGACGAGGACATGATCAGTCAAGCATATGAGCTGTCC 1980  
| | | | |  
Db 15820 GGTCTCAGCTGTGTGTGACACGACGAGGACATGATCAGTCAAGCATATGAGCTGTCC 15879  
| | | | |  
QY 1981 ACACTGAAAGACGTGTCTCAACAGCCTGAATCAATGAGTTAGCTTAATGAATAAAT 2040  
| | | | |  
Db 15880 ACACTGAAAGACGTGTCTCAACAGCCTGAATCAATGAGTTAGCTTAATGAATAAAT 15939  
| | | | |  
QY 2041 CCCAGACTACTTACGCTTTTAATGCTTTTATTCATAAAACGTGTGAAGACTGACGAA 2100  
| | | | |  
Db 15940 CCCAGACTACTTACGCTTTTAATGCTTTTATTCATAAAACGTGTGAAGACTGACGAA 15999  
| | | | |  
QY 2101 CCATTGGAACATTTAACTCAGACTGTGATTCGAGATTCGGGAACTTTAGTTCTATCTG 2160  
| | | | |

Db 1600 CCATTGGAACATTTAACTCAGACTGTGATTCAGAGTGGGAAACCTTATGTTCTATCTG 16059  
| | | | |  
QY 2161 AATCCAGAGACGCACACCTTACTATATCTGCCCAACTAATGATTAATTAATAC 2216  
| | | | |  
Db 16060 AATCCAGAGACGCACACCTTACTATATCTGCCCAACTAATGATTAATTAATTAATAC 16115  
| | | | |  
RESULT 2  
ABAI4559  
ID ABAI4559 standard; DNA; 16650 BP.  
XX  
XX ABAI4559;  
XX  
XX 23-JAN-2002 (first entry)  
XX  
XX Human nervous system related polynucleotide SEQ ID NO 6890.  
XX  
XX Human; nootropic; neuroprotective; cyostatic; dermatological; viroicide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnereary;  
KW antiparkinsonian; antislaking; antihaemic; antiarthritic; cancer;  
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;  
KW antiallergic; antidiabetic; antilicer; anticonvulsant; antifungal;  
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; de.  
OS Homo sapiens.  
XX  
XX WO200159063-A2.  
XX  
XX 16-AUG-2001.  
XX  
XX 17-JAN-2001; 2001MO-US001334.  
PF  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214866P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218390P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225470P.  
PR 14-AUG-2000; 2000US-0225471P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.

[illegible]

Query Match	Best Local Similarity	Matches 2205; Conservative	Score 98.8%; Pred. 99.5%;	Mismatches 0;	Indels 10;	Gaps 1;	DB 5;	Length 16650;	Other;
1	ATCTCTGTTTCTTCTCCAGCCTGAACAAAGAGGAGTGGAGTCCAGGCCCAATCAGAGG	60							
11879	ATCTCTGTTTCTTCTCCAGCCTGAACAAAGAGGAGTGGAGTCCAGGCCCAATCAGAGG	11936							
61	ACGAGGAGGATGCTCAGCGCTGGCATATTGGCCGCTTCTTAGGCCCTGCAGTCCCATGATC	120							
11939	ACGAGGAGGATGCTCAGCGCTGGCATATTGGCCGCTTCTTAGGCCCTGCAGTCCCATGATC	11996							
121	ACCATCTCCACGAGGCAACTCAGAGCAAACTATTGTGGCCCAAGAGGAGTCCCTTC	180							
11999	ACCATCTCCACGAGGCAACTCAGAGCAAACTATTGTGGCCCAAGAGGAGTCCCTTC	180							

QY 181 TCCAGAGGCGCTGCCCCCAACAGAGGAGCCCAAGAGCTTAGGGGCGAGAG 240  
Db 12059 TCCAGAGGCGCTGCCCCCAACAGAGGAGCCCAAGAGCTTAGGGGCGAGAG 12118  
QY 241 ACAACAGGCGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATGAGA 300  
Db 12119 ACAACAGGCGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCCCACTGTATGAGA 12178  
QY 301 GGGCAGGCTACTACAGTGGCCCAACAGAGGCGCCCTCAGCCCACTCCCAAGTCTTCCCCC 360  
Db 12179 GGGCAGGCTACTACAGTGGCCCAACAGAGGCGCCCTCAGCCCACTCCCAAGTCTTCCCCC 12238  
QY 361 TAGAACCATGACGCGCTCAAGAGCTTCAAGTGTCAAGGCAATGAGACCAAGAGAGAAA 420  
Db 12239 TAGAACCATGACGCGCTCAAGAGCTTCAAGTGTCAAGGCAATGAGACCAAGAGAGAAA 12298  
QY 421 GCTTAAAGACTGTGAGTCTTGCGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 480  
Db 12299 GCTTAAAGACTGTGAGTCTTGCGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 12358  
QY 481 GGGCCTTGATGAGAGCAGCCAGAGATATCTCAAGTGAAGAGAGAAAAGTGTGAATTAAC 540  
Db 12359 GGGCCTTGATGAGAGCAGCCAGAGATATCTCAAGTGAAGAGAGAAAAGTGTGAATTAAC 12418  
QY 541 TGAAGGATATGAGAGATCCCGAAGATCACTCAAGAACTTTTGAACATCAACCA 600  
Db 12419 TGAAGGATATGAGAGATCCCGAAGATCACTCAAGAACTTTTGAACATCAACCA 12478  
QY 601 CCAACATACACATCACTCAAGATCACTGAATGATGATCTTATTTGGGCTTTGAAAAACAGT 660  
Db 12479 CCAACATACACATCACTCAAGATCACTGAATGATGATCTTATTTGGGCTTTGAAAAACAGT 12538  
QY 661 CTGTCTCCCACTGGAACCAAGGGGCACTGCAATGCGCTGTGCCCCCACTCCAGCTTCCCTTG 720  
Db 12539 CTGTCTCCCACTGGAACCAAGGGGCACTGCAATGCGCTGTGCCCCCACTCCAGCTTCCCTTG 12598  
QY 721 CTGTAGAGCTACCTTCTCTCACAAATTTCTAGGGGTTCCATCACTGTGCAAGAGCAGACTGG 780  
Db 12599 CTGTAGAGCTACCTTCTCTCACAAATTTCTAGGGGTTCCATCACTGTGCAAGAGCAGACTGG 12658  
QY 781 ACCTAGCGCCAGCAGCTGCGCTTGGGGTATATCTTGGCCACCTTTCACAGAGATCTAGGGA 840  
Db 12659 ACCTAGCGCCAGCAGCTGCGCTTGGGGTATATCTTGGCCACCTTTCACAGAGATCTAGGGA 12718  
QY 841 AGTGTTCGGGACCTTTTCTCACTTCACTCCGCTGTATACCCGGAAGACTTCTTGGAGCAG 900  
Db 12719 AGTGTTCGGGACCTTTTCTCACTTCACTCCGCTGTATACCCGGAAGACTTCTTGGAGCAG 12778  
QY 901 GTGAAGGAAGATGAGTGTGTGCTGACCAAGATGCTGTGGAAGACTGCCCCAGGGGCTGAC 960  
Db 12779 GTGAAGGAAGATGAGTGTGTGCTGACCAAGATGCTGTGGAAGACTGCCCCAGGGGCTGAC 12838  
QY 961 AGGCGAGGCTTAGCTGAGCAGATGTATCACTGCGCCCACTTATTTGAAGAGGGTGG 1020  
Db 12839 AGGCGAGGCTTAGCTGAGCAGATGTATCACTGCGCCCACTTATTTGAAGAGGGTGG 12898  
QY 1021 CTGACCCCAAAACATGAGGTGGGAGTCACTGATGACAGATGAACCTTCCCCCATTAAC 1080  
Db 12899 CTGACCCCAAAACATGAGGTGGGAGTCACTGATGACAGATGAACCTTCCCCCATTAAC 12958  
QY 1081 TATTTAGGAGTAGTACCAAGACATCAAGAAAGGGTGGAGAACTGCTCACTCCCTAGG 1140  
Db 12959 TATTTAGGAGTAGTACCAAGACATCAAGAAAGGGTGGAGAACTGCTCACTCCCTAGG 13018  
QY 1141 AACTGTAGATGTGTGAGGTGTGAGGGGTGTCCAGCGCCCTTAGTATTTTCTCACTGCTG 1200  
Db 13019 AACTGTAGATGTGTGAGGTGTGAGGGGTGTCCAGCGCCCTTAGTATTTTCTCACTGCTG 13078  
QY 1201 GGAACCTCAGCAAAATCTTCTTCTCTTGGGGGTGAGCCCAAGAGTGTCACAAAATCA 1260  
Db 13079 GGAACCTCAGCAAAATCTTCTTCTCTTGGGGGTGAGCCCAAGAGTGTCACAAAATCA 13138  
QY 1261 GATATTTTCCCTTTATTCAGATTTTCTGAGACATGTCAACCAATTAATAACCCCACTT 1320

Db 13139 GATATTTTCCCTTTATTCAGATTTTCTGAGACATTTCAACCAATTAATAACCCCACTT 13198  
QY 1321 CAGCCCAATCAGGTGGAGAGAGTGAATCTTCCCTTTTGTGATTTCTCAAGAGTAACT 1380  
Db 13199 CAGCCCAATCAGGTGGAGAGAGTGAATCTTCCCTTTTGTGATTTCTCAAGAGTAACT 13258  
QY 1381 TTCAAGGGTCAAGAACGAGCTATTATGATTTGAACCTTAATAAGGCAACATTTGANT 1440  
Db 13259 TTCAAGGGTCAAGAACGAGCTATTATGATTTGAACCTTAATAAGGCAACATTTGANT 13317  
QY 1441 CTGTCTTGAAGCTTGAAGACAGAACTTGGCAAACTGTGCGCTGTTCAGCAAGAGATG 1500  
Db 13318 CTGTCTTGAAGCTTGAAGACAGAACTTGGCAAACTGTGCGCTGTTCAGCAAGAGATG 13377  
QY 1501 TTCAATTTAAGAAATGTGTCTTGGGCTGGGTGTGAGGCAATGAAATCAAGAGGTCA 1560  
Db 13378 TTCAATTTAAGAAATGTGTCTTGGGCTGGGTGTGAGGCAATGAAATCAAGAGGTCA 13437  
QY 1561 GGAAGTTGAGACCACTGGCCCAACATGATGAACCCCATCTCTACCAAAAAAATACAA 1620  
Db 13438 GGAAGTTGAGACCACTGGCCCAACATGATGAACCCCATCTCTACCAAAAAAATACAA 13497  
QY 1621 ATCAGTGGCGCTGTGTGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGAGAAATGTC 1680  
Db 13498 ATCAGTGGCGCTGTGTGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGAGAAATGTC 13557  
QY 1681 TTGAACCCAGAGAGGT 1740  
Db 13558 TTGAACCCAGAGAGGT 13617  
QY 1741 GGAAGTGAAGCTGTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800  
Db 13618 GGAAGTGAAGCTGTCTCAAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 13677  
QY 1801 GCAACATTTTGTATTTGAAGATGAAGATGATCTTCACTTATGCTGTATTAATCTTTCAT 1860  
Db 13678 GCAACATTTTGTATTTGAAGATGAAGATGATCTTCACTTATGCTGTATTAATCTTTCAT 13737  
QY 1861 TCTCAGTGAAGATGAAGACATTTCTTCACTGCTTCTTCAATGGGAGTCTTGGCAGCA 1920  
Db 13738 TCTCAGTGAAGATGAAGACATTTCTTCACTGCTTCTTCAATGGGAGTCTTGGCAGCA 13797  
QY 1921 GGTCTCAGCTGT 1980  
Db 13798 GGTCTCAGCTGT 13857  
QY 1981 ACACTGAAGACGTGTCTTCAACAGCAGCTGAATCAATGTGTAGTTAATGAATTAAT 2040  
Db 13858 ACACTGAAGACGTGTCTTCAACAGCAGCTGAATCAATGTGTAGTTAATGAATTAAT 13917  
QY 2041 CCCAGACTACTGAGCTTTAATGCTTTTATTAATTAATAAAGCTGAGACTGAA 2100  
Db 13918 CCCAGACTACTGAGCTTTAATGCTTTTATTAATTAATAAAGCTGAGACTGAA 13977  
QY 2101 CCAATGGAAACATTTAATCAAGCTGTGAATTTAGAGTGGGGAACCTTAGTCTTATCTG 2160  
Db 13978 CCAATGGAAACATTTAATCAAGCTGTGAATTTAGAGTGGGGAACCTTAGTCTTATCTG 14037  
QY 2161 AATCAAGACAGCAGCACTTATGATATGCCCCAATCAATGATTAATAATATAC 2216  
Db 14038 AATCAAGACAGCAGCACTTATGATATGCCCCAATCAATGATTAATAATATAC 14093

RESULT 3  
ABAI4557  
ID ABAI4557 strand; DNA; 18530 BP.  
XX ABAI4557;  
XX  
XX 23-JAN-2002 (first entry)  
XX Human nervous system related polynucleotide SEQ ID NO 6886.



Query Match	Best Local Similarity	Matches 2205	Conservative	Score 2188.6	DB 5	Length 18530	Mismatches 0	Pred. No. 0	Indels 10	Gaps 1	Other
0Y	1	ATCTCGTGTCTCTTCCAGCCTTGAA	98.8%	99.5%	0	10	1	0	0	0	0
DB	11879	ATCTCTGTTCTTCTTCCAGCTGAA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	61	ACGAGAGAGATGCTCAGCGTGGCAT	98.8%	99.5%	0	10	1	0	0	0	0
DB	11939	ACGAGAGAGATGCTCAGCGTGGCAT	98.8%	99.5%	0	10	1	0	0	0	0
0Y	121	ACCATCTCTCCCAAGGCAAACTCA	98.8%	99.5%	0	10	1	0	0	0	0
DB	11999	ACCATCTCTCCCAAGGCAAACTCA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	181	TCGACGAGGGGCTGCTCCCAAAA	98.8%	99.5%	0	10	1	0	0	0	0
DB	12059	TCGACGAGGGGCTGCTCCCAAAA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	241	ACAACAAGGCGCTTGAGCTTAA	98.8%	99.5%	0	10	1	0	0	0	0
DB	12119	ACAACAAGGCGCTTGAGCTTAA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	301	GGCCAGGCTACTAAGTGGCCCA	98.8%	99.5%	0	10	1	0	0	0	0
DB	12179	GGCCAGGCTACTAAGTGGCCCA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	361	TAGAAACCATCAGGCGCGTCA	98.8%	99.5%	0	10	1	0	0	0	0
DB	12239	TAGAAACCATCAGGCGCGTCA	98.8%	99.5%	0	10	1	0	0	0	0
0Y	421	GCTTAAAGACTGTGATGTTCTG	98.8%	99.5%	0	10	1	0	0	0	0

Db	12299	GCTTAAAGACTGTAGTCTGGGGCCAGAAAAGTTTGAATTGCTCTCAGAGCGATG	12358
Oy	481	GGGCGCTTGAATGAGACCCAGAAAGTATCTCAAGTGAGAGGAAAACCTGAGAGTTAAC	540
Db	12359	GGGCGCTTGAATGAGACCCAGAAAGTATCTCAATGAGAGGAAAACCTGAGAGTTAAC	12418
Oy	541	TGACGGAATGCGCAGAGATCCCGAAAATCACTCAAGAACCTTTGGAAACAATCCAA	600
Db	12419	TGACGGAATGCGCAGAGATCCCGAAAATCACTCAAGAACCTTTGGAAACAATCCAA	12478
Oy	601	CCAAATATACATCACTCAAGATACATGATGATCTTATTTGGGCGTTGGAACAGAT	660
Db	12479	CCAAATATACATCACTCAAGATACATGATGATCTTATTTGGGCGTTGGAACAGAT	12538
Oy	661	CTGTCTCTCAACTGAAACGAGGGGCACTGACCTTCCGTGGGCCCAACCCAGCTTCCCTTG	720
Db	12539	CTGTCTCTCAACTGAAACGAGGGGCACTGACCTTCCGTGGGCCCAACCCAGCTTCCCTTG	12598
Oy	721	CTGTAGCCCTAACCTTCTCTCACAATTTCTTGAAGGTTTCATCACTGCGCAGACACCTGG	780
Db	12599	CTGTAGCCCTAACCTTCTCTCACAATTTCTTGAAGGTTTCATCACTGCGCAGACACCTGG	12658
Oy	781	ACCTAACCCCAACGACCTGGGCTTTGGGGATATATCTTTGGCCACTTCAACAGGGAATCTTAGGA	840
Db	12659	ACCTAACCCCAACGACCTGGGCTTTGGGGATATATCTTTGGCCACTTCAACAGGGAATCTTAGGA	12718
Oy	841	AGTGTGGGGACCTTTTCTCACTTCAACCCCTGGATACCCCGGAAGACTTTTGGGACGAG	900
Db	12719	AGTGTGGGGACCTTTTCTCACTTCAACCCCTGGATACCCCGGAAGACTTTTGGGACGAG	12778
Oy	901	GTAAGGAAGATGAGGTTGTGTGACCAAGATCTGTGGAAGACCTGGCCCAAGGCTGAC	960
Db	12779	GTAAGGAAGATGAGGTTGTGTGACCAAGATCTGTGGAAGACCTGGCCCAAGGCTGAC	12838
Oy	961	AGGCGAGGCTTATGCTGAGCAGATTTATCATCTGGCCCCCACTTATCTTTAGCAAGGATGG	1020
Db	12839	AGGCGAGGCTTATGCTGAGCAGATTTATCATCTGGCCCCCACTTATCTTTAGCAAGGATGG	12898
Oy	1021	CTGACCCCAAAACCATGAGGTGGGACGTGAGATGACAGATGAAACATCTCCGCCATAC	1080
Db	12899	CTGACCCCAAAACCATGAGGTGGGACGTGAGATGACAGATGAAACATCTCCGCCATAC	12958
Oy	1081	TATTTAGGGATGATCCCAACCATGAGAAAGGATGGAGCACTGCTCTAG	1140
Db	12959	TATTTAGGGATGATCCCAACCATGAGAAAGGATGGAGCACTGCTCTAG	13018
Oy	1141	AACTGTAGATGCTGAGGTTGAGGGGTGTCCAGCGCCCTTAGTCAATTTTCTCACTGCTG	1200
Db	13019	AACTGTAGATGCTGAGGTTGAGGGGTGTCCAGCGCCCTTAGTCAATTTTCTCACTGCTG	13078
Oy	1201	GGAAACCTCACAAAATCTTCTGCTCTCTTGGGGGTCAACCCAAAGCTGTCAAAAATCA	1260
Db	13079	GGAAACCTCACAAAATCTTCTGCTCTCTTGGGGGTCAACCCAAAGCTGTCAAAAATCA	13138
Oy	1261	GATATTTCCCTTATTTCCAGATTTCCCTGGACACTGTCAACCAATTTAACAACCCCACTT	1320
Db	13139	GATATTTCCCTTATTTCCAGATTTCCCTGGACACTGTCAACCAATTTAACAACCCCACTT	13198
Oy	1321	CAGCCCAATCAAGTGGAGGAAGTGTAACTTCCCTTTCTGGAATCTCAACAGTACT	1380
Db	13199	CAGCCCAATCAAGTGGAGGAAGTGTAACTTCCCTTTCTGGAATCTCAACAGTACT	13258
Oy	1381	TTTCAAGGATCAACAGCATATTTATGATTTGAACCTTTAAAGGGCAAAATTTCAAT	1440
Db	13359	TTTCAAGGATCAACAGCATATTTATGATTTGAACCTTTAAAGGGCAAAATTTCAAT	13317
Oy	1441	CTTGTCTCTAGAGGCTAGACAGAACTTTGGCAAACTGTGGCTGTTCAGCAAAAGATG	1500
Db	13318	CTTGTCTCTAGAGGCTAGACAGAACTTTGGCAAACTGTGGCTGTTCAGCAAAAGATG	13377
Oy	1501	TTTCAATTTAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAAGAGGTCA	1560

[illegible][illegible]

PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239355P.  
PR 13-OCT-2000; 2000US-0239357P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249301P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 01-DEC-2000; 2000US-0251160P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX  
DR MPI; 2001-541565/60.  
XX  
XX  
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system cancers  
and metastases.

XX  
PS Disclosure; SEQ ID NO 6889; 1701bp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins  
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (anti)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing  
CC; (e) neurological diseases such as cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;  
Query Match 98.7%; Score 2187; DB 5; Length 18537;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
QY 1 ATCTCTGTTCTTTTCCAGCTGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGG 60  
DB 11885 ATCTCTGTTCTTTTCCAGCTGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGG 11944  
QY 61 ACAGAGAGATGTCACGCTGAGATGATGCGGCTTCTTGAAGCTGCAATGCCATGATC 120  
DB 11945 ACAGAGAGATGTCACGCTGAGATGATGCGGCTTCTTGAAGCTGCAATGCCATGATC 12004  
QY 121 ACCATCTCCAGGGGAACTCAAGACCAAACTACTGTGCCCCAAGAGGAAATCCCTTC 180  
DB 12005 ACCATCTCCAGGGGAACTCAAGACCAAACTACTGTGCCCCAAGAGGAAATCCCTTC 12064  
QY 181 TCACAGAGGGCTGCCCCAAAACCAAGAGCAGCCAAAGAAAGTTAGGGGCCAGAGAG 240  
DB 12065 TCACAGAGGGCTGCCCCAAAACCAAGAGCAGCCAAAGAAAGTTAGGGGCCAGAGAG 12124  
QY 241 ACAACAAGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCTGATTCAGA 300  
DB 12125 ACAACAAGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCTGATTCAGA 12184  
QY 301 GGCCAGGCTACTACAGTGCAGCCACAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCCC 360  
DB 12185 GGCCAGGCTACTACAGTGCAGCCACAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCCC 12244  
QY 361 TAGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCATAGACCAAAAGACAAA 420  
DB 12245 TAGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGCATAGACCAAAAGACAAA 12304  
QY 421 GCTTAAGAAGTGTGAATTTCTGGGGCCAAAGAAAGTTTGAATGTCTTCAGAGAGGATG 480  
DB 12305 GCTTAAGAAGTGTGAATTTCTGGGGCCAAAGAAAGTTTGAATGTCTTCAGAGAGGATG 12364  
QY 481 GGCCCTTGATGAGCAACCAAGAGTATCTCAAGTGTGAGAGAAAGCTGTGAGATTAAAC 540  
DB 12365 GGCCCTTGATGAGCAACCAAGAGTATCTCAAGTGTGAGAGAAAGCTGTGAGATTAAAC 12424  
QY 541 TGACGGATATGCCAGAGATCCCGAAATATCACTCAAGAAAGCTTTGAAACATACCAA 600  
DB 12425 TGACGGATATGCCAGAGATCCCGAAATATCACTCAAGAAAGCTTTGAAACATACCAA 12484  
QY 601 CCAACATACACTACACTCAAAAGATCAATGATATCTTATTTGGGCTTGGAACAAAGT 660  
DB 12485 CCAACATACACTACACTCAAAAGATCAATGATATCTTATTTGGGCTTGGAACAAAGT 12544  
QY 661 CTGTCTCCAGCTGAACCAAGGGGCACTGATGCGCTGTGCCCCAGCTTCCCTTG 720



Db	12545	CTGTCCTCCACCTGAACCAAGGAGCACTGATTTGGCCCTGTGCCCCACCCACGCTTCCCTTG	12604
QY	721	CTGTGAGCTTACCCCTTCTCCACAATTTCCATAGGGTTCCATCATCTGCACAGACACTGG	780
Db	12605	CTGTGAGCTTACCCCTTCTCCACAATTTCCATAGGGTTCCATCATCTGCACAGACACTGG	12664
QY	781	ACCTAAGCCCAAGCACTGGCTTTGGGCTTATATCTTTGGCACTTCAAGGAGTCTTAGGGA	840
Db	12665	ACCTAAGCCCAAGCACTGGCTTTGGGCTTATATCTTTGGCACTTCAAGGAGTCTTAGGGA	12724
QY	841	AGTGTGGGAGCCCTTTCTCACTTCAACCCCGGATATACCCGGAAAGCTTTTGGGACAG	900
Db	12725	AGTGTGGGAGCCCTTTCTCACTTCAACCCCGGATATACCCGGAAAGCTTTTGGGACAG	12784
QY	901	GTTAAAGGAAGATGAGGTGTGTCTGACCAAGATCTGTGTGAGAACCTGCCACAGGGCTTAC	960
Db	12785	GTTAAAGGAAGATGAGGTGTGTCTGACCAAGATCTGTGTGAGAACCTGCCACAGGGCTTAC	12844
QY	961	AGGCCAGGCTTACCTGAGCAGATGTATATCACTAGGCCCAACTTACTTTGACCAAGGTTGG	1020
Db	12845	AGGCCAGGCTTACCTGAGCAGATGTATATCACTAGGCCCAACTTACTTTGACCAAGGTTGG	12904
QY	1021	CTGACCCCAACCAATGAGGTGGCAGTCAAGCTGATGACAGATGAACTTCCCATTAC	1080
Db	12905	CTGACCCCAACCAATGAGGTGGCAGTCAAGCTGATGACAGATGAACTTCCCATTAC	12964
QY	1081	TATTTAGGTTAGTACCAAGACATPACAGAAAGGTTGCAGGAACTGCTCATCTCTAAG	1140
Db	12965	TATTTAGGTTAGTACCAAGACATPACAGAAAGGTTGCAGGAACTGCTCATCTCTAAG	13024
QY	1141	AACGTGTAGTGTGAGGTTGAGGGTGTCAAGCCCTTATAGTATTTTCTCATCTGCTG	1200
Db	13025	AACGTGTAGTGTGAGGTTGAGGGTGTCAAGCCCTTATAGTATTTTCTCATCTGCTG	13084
QY	1201	GGAACTTCACCAAAATATCTTCTCTCTCTTGGGGTACGCCAAAGCTGTCAAAATTA	1260
Db	13085	GGAACTTCACCAAAATATCTTCTCTCTCTTGGGGTACGCCAAAGCTGTCAAAATTA	13144
QY	1261	GATATTTCCCTTTATTTCCAGATTTCTGTGACACTGTCAACCATTATTAACAACCCCAT	1320
Db	13145	GATATTTCCCTTTATTTCCAGATTTCTGTGACACTGTCAACCATTATTAACAACCCCAT	13204
QY	1321	CAGCCCAATCACTGGGAGAGGTGTAATTCCTTCTGTGATTTCTCAGCAGTACT	1380
Db	13205	CAGCCCAATCACTGGGAGAGGTGTAATTCCTTCTGTGATTTCTCAGCAGTACT	13264
QY	1381	TTGACGGGTACAGAACCGCAGCTATTTATGTAATGAAACCTTAAAGGCAACAATTCANT	1440
Db	13265	TTGAC -GGTACAGAACCGCAGCTATTTATGTAATGAAACCTTAAAGGCAACAATTCANT	13323
QY	1441	CTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1500
Db	13324	CTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	13383
QY	1501	TTTCATATTTTAAAGATCTTGTCTTGGGCTGGGTGTGTGAGGCAAGTAAATCACAGAGGTCA	1560
Db	13384	TTTCATATTTTAAAGATCTTGTCTTGGGCTGGGTGTGTGAGGCAAGTAAATCACAGAGGTCA	13443
QY	1561	GGAGTTTGAGCACTGCGCAACATATGAAACCCCATCTTCTTCAAAAATAATACA	1620
Db	13444	GGAGTTTGAGCACTGCGCAACATATGAAACCCCATCTTCTTCAAAAATAATACA	13503
QY	1621	ATGAGCTGGCGGTGTGTGTGTCTGTGTAGTCCCAAGCAGAGGTTTGAAGGGAGAAATTCG	1680
Db	13504	ATGAGCTGGCGGTGTGTGTGTCTGTGTAGTCCCAAGCAGAGGTTTGAAGGGAGAAATTCG	13563
QY	1681	TTGAACCCAGAGGTGTGTGTGTGAAGGATTTGAGCACTGCAATTCAGCTGGGCGAC	1740
Db	13564	TTGAACCCAGAGGTGTGTGTGTGAAGGATTTGAGCACTGCAATTCAGCTGGGCGAC	13623
QY	1741	GGAGTGTGACTGTCTCAAAAATAATGATGTCTCAACTTTTGGCCCTCTTACT	1800
Db	13624	GGAGTGTGACTGTCTCAAAAATAATGATGTCTCAACTTTTGGCCCTCTTACT	13683

Qy	1801	GCACATTTTGGTATTTGAATGAAGGTACCTTCCATTAATTATGCTGTAAATACCTTTCAT	1860
Dp	13864	GCAACATTTTGGTATTTGAATGAAGGTACCTTCCATTAATTATGCTGTAAATACCTTTCAT	13744
Qy	1861	TCTCAGTAGGGATGAAGCACAATTCCTAACCTGTTCTCTAATGGGGATGCTTGGCAGCCA	1920
Dp	13744	TCTCAGTAGGGATGAAGCACAATTCCTAACCTGTTCTCTAATGGGGATGCTTGGCAGCCA	13803
Qy	1921	GGTCTCTACCTGTGTGTACACGACGACGACACTGTATCCAGTCAGACGCCATAAGCTGTCC	1980
Dp	13804	GGTCTCTACCTGTGTGTACACGACGACGACACTGTATCCAGTCAGACGCCATAAGCTGTCC	13863
Qy	1981	ACACTGAANAAGGTGTCTTCACAAACAGCCTGAATCAATAGTTAGCTTAATAGATAAAT	2040
Dp	13864	AACACTGAANAAGGTGTCTTCACAAACAGCCTGAATCAATAGTTAGCTTAATAGATAAAT	13923
Qy	2041	CCGACACTACTCAGCCTTTAATGCTTTTATTCATAATAAACTGTGAAAGCTAGACTGAA	2100
Dp	13924	CCGACACTACTCAGCCTTTAATGCTTTTATTCATAATAAACTGTGAAAGCTAGACTGAA	13983
Qy	2101	CCATTGAAAACATTAACTCAGACTCTGGATTACAGATCGGGAACCTTAAGTTCTATCTG	2160
Dp	13984	CCATTGAAAACATTAACTCAGACTCTGGATTACAGATCGGGAACCTTAAGTTCTATCTG	14043
Qy	2161	AATCCAGAAGCGACACCTTAGTATCTGCCCAAATTAATAGTTAATTAATATAC	2216
Dp	14044	AATCCAGAAGCGACACCTTAGTATCTGCCCAAATTAATAGTTAATTAATATAC	14099

## RESULT 5

ABL8969 / standard; cDNA; 1717 BP,  
ID  
XX

AC ABL89697;

DI 24-MAY-2002 (first entry)  
XX

Human polynucleotide SEQ ID NO 259.

KW cytotoxic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotrophic; antidabetic; antiinflammatory; anticancer;  
KW vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW caridant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; gene; ss.

OS Homo sapiens.  
YY

PN WO200190304-A2  
XX

PD 29-NOV-2001  
XX

18-MAY-2001; 2001WO-US016450.

19-MAY-2000; 2000US-0205515P.  
FK  
XX

PA (HUMA-) HUMAN GENOME SCI INC.  
XX

PLISE CE, Rosen CA,  
XX

DR WP1; 2002-122018/16.  
DR P-PSDB: ABB90399

Novel 1405 1807 24

PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders.

PS Claim 4; SEQ ID NO 259; 2081pp + Sequence Listing; English  
XX

the literature relates to novel genes (ABL98449-ABL90853) and proteins (ABB890440-ABB90444) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are



CC Isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in  
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and  
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,  
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing  
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)  
CC infectious diseases such as viral, bacterial, fungal and parasitic  
CC infections. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [http://wipo.int/pub/published\\_pcr\\_sequences](http://wipo.int/pub/published_pcr_sequences)

XX Sequence 1717 BP; 464 A; 460 C; 417 G; 367 T; 0 U; 9 Other;

Query Match 73.6%; Score 1631.6; DB 6; Length 1717;

Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1692; Conservative 3; Mismatches 18; Indels 5; Gaps 5;

```
QY 57 GAGACGAGAGAGATGCTCAGCGTCGATCATTTGGCGCTTCTTGAAGCTTCCAT 116
DB 1 GAGACGAGAGAGATGCTCAGCGTCGATCATTTGGCGCTTCTTGAAGCTTCCAT 60
QY 117 GATCACCATCTCCAGGGGAACTCAAGGACCAACTAGTGGGCGCCAGAGGAATCC 176
DB 61 GATCACCATCTCCAGGGGAACTCAAGGACCAACTAGTGGGCGCCAGAGGAATCC 120
QY 177 CTCTCCAGAGGGGCTGCCCAAAAACCAAGGACCCAAAGAAAGTTAAGGGGCGAG 236
DB 121 CTCTCCAGAGGGGCTGCCCAAAAACCAAGGACCCAAAGAAAGTTAAGGGGCGAG 179
QY 237 GAAAGCAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCATGTAT 236
DB 180 GAAAGCAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCATGTAT 239
QY 297 CAGAGGCGAGGCTACTCAGGCGCCCAAGAGCGCCCTCAGCCCATCTCCAGTCTTTC 356
DB 240 CAGAGGCGAGGCTACTCAGGCGCCCAAGAGCGCCCTCAGCCCATCTCCAGTCTTTC 299
QY 357 CCCTAGAACCATCAGCGCGCTCAAGCTTCAAGTGTCAAGGATAGACCAAGAAC 416
DB 300 CCCTAGAACCATCAGCGCGCTCAAGCTTCAAGTGTCAAGGATAGACCAAGAAC 359
QY 417 AAAAGCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTTGAATTGCTTCAGAGAGC 476
DB 360 AAAAGCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTTGAATTGCTTCAGAGAGC 419
QY 477 GATGGGGCTTGTATGAGGACCCAGAGATATCTCAATGAGAGAGAAACTGTGAGATT 536
DB 420 GATGGGGCTTGTATGAGGACCCAGAGATATCTCAATGAGAGAGAAACTGTGAGATT 479
QY 537 AACCTGACGATATGACAGAGATCCCGAAATCACTCAAGAAAGCTTTGGAACAATCA 596
DB 480 AACCTGACGATATGACAGAGATCCCGAAATCACTCAAGAAAGCTTTGGAACAATCA 539
QY 597 CCAACCAACATACACTCACTCAATCAATGATCAATGATCTTATTTGGGCTTTGAAAC 656
DB 540 CCAACCAACATACACTCACTCAATCAATGATCAATGATCTTATTTGGGCTTTGAAAC 599
QY 657 AGGTCTGTCTCTCAACCTGAACCAAGGGGCACTGATGGCGCCGACCCCAAGCTTCC 716
DB 600 AGGTCTGTCTCTCAACCTGAACCAAGGGGCACTGATGGCGCCGACCCCAAGCTTCC 659
QY 717 CTGTCTGTGAGCTTACCTTCTCAACAATTTCTTGAAGGTTCCATCACTGACAGAGACA 776
DB 660 CTGTCTGTGAGCTTACCTTCTCTCAACAATTTCTTGAAGGTTCCATCACTGACAGAGACA 719
QY 777 CTGAGACTTACGCCAGACCTTGGGCTTGGGATATATCTTGGCCACCTTACAGAGGATCTTA 836
DB 720 CTGAGACTTACGCCAGACCTTGGGCTTGGGATATATCTTGGCCACCTTACAGAGGATCTTA 779
```

```
QY 837 GGGAGATGTCGGGACCTTTTCTCACTTCACTCCCTGTATGACCCGGAGAACTTCTTGGA 896
DB 780 GGGAGATGTCGGGACCTTTTCTCACTTCACTCCCTGTATGACCCGGAGAACTTCTTGGA 839
QY 897 CCAAGTGAAGAAAGATGAGGTTGTGCTGACAGAAATGCTGTGAGAACTGCCCGAGGC 956
DB 840 CCAAGTGAAGAAAGATGAGGTTGTGCTGACAGAAATGCTGTGAGAACTGCCCGAGGC 899
QY 957 TGACAGGCGAGGCTTGTAGTGTAGAGATGATTTATCATGCGCCCACTTATTTGAGCAAG 1016
DB 900 TGACAGGCGAGGCTTGTAGTGTAGAGATGATTTATCATGCGCCCACTTATTTGAGCAAG 959
QY 1017 GTGAGTGAACCAAAACCATGAGGTCAGTCACTGATGATGATGATGATGATGATGATGATG 1076
DB 960 GTGAGTGAACCAAAACCATGAGGTCAGTCACTGATGATGATGATGATGATGATGATGATG 1019
QY 1077 TAACTATTTAGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1136
DB 1020 TAACTATTTAGGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
QY 1137 TAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1196
DB 1079 TAGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1138
QY 1197 CTTGGGAACTGACCAAAATACCTTCTTCTTCTTGGGGGTCAGGCCAAAGCTGTCAACAA 1256
DB 1139 CTTGGGAACTGACCAAAATACCTTCTTCTTCTTGGGGGTCAGGCCAAAGCTGTCAACAA 1198
QY 1257 ATCAGATATTTCCCTTATTCAGATTTCTGTGACACTGTACCCCAATTTAAACACCC 1316
DB 1199 ATCAGATATTTCCCTTATTCAGATTTCTGTGACACTGTACCCCAATTTAAACACCC 1258
QY 1317 ACTTACGCCCAATCAGTGTGAGAGAAATGTAATCTTCTTCTTGTGATTTCTCAGAGCT 1376
DB 1259 ACTTACGCCCAATCAGTGTGAGAGAAATGTAATCTTCTTCTTGTGATTTCTCAGAGCT 1318
QY 1377 TACTTTCAGGGGTCAGAACGAGCTATTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1436
DB 1319 TACTTTCAGGGGTCAGAACGAGCTATTAATGTAATGTAATGTAATGTAATGTAATGTAATG 1377
QY 1437 CANTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1496
DB 1378 CANTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1437
QY 1497 GATGTTCAATTTTAAAGATTTTGTCTTGGGCTGTGAGGCAAGTGAATACAGAGAG 1556
DB 1438 GATGTTCAATTTTAAAGATTTTGTCTTGGGCTGTGAGGCAAGTGAATACAGAGAG 1497
QY 1557 GTGAGAGTTTGAACCAACCTGCGCAACATGATGAAACCCCATCTTACCAAAAAAAT 1616
DB 1498 GTGAGAGTTTGAACCAACCTGCGCAACATGATGAAACCCCATCTTACCAAAAAAAT 1557
QY 1617 ACAAAATCAGCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1676
DB 1558 ACAAAATCAGCTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1617
QY 1677 TTGCTTGAACCCAGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1735
DB 1618 TTGCTTGAACCCAGAGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1677
QY 1736 GCGAGC-GGAGTGAAGTGTCTCAAAAAAAGAAAAA 1772
DB 1678 GCGAGCAGAGTGAAGTGTCTCAAAAAAAGAAAAA 1715
```

RESULT 6

AAV99722 standard; cDNA; 1263 BP.

AAV99722;

26-APR-1999 (first entry)





Db 1442 GCTTAAGGCTGTGAGCGGCTTCAAGTCTGCGCCCACTGATCATAGAGCCAGGCTAATAAG 1501  
Qy 317 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 376  
Db 1502 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 1561  
Qy 377 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAAGCTTAAAGCTGTAG 436  
Db 1562 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAAGCTTAAAGCTGTAG 1621  
Qy 437 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGGATGGGCTTATGAGCA 496  
Db 1622 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGGATGGGCTTATGAGCA 1681  
Qy 497 CCCAGAAGTATCTCAAGTGAAGGAAAGCTTGAAGCTTAACTGAGATATGCGAGA 556  
Db 1682 CCCAGAAGTATCTCAAGTGAAGGAAAGCTTGAAGCTTAACTGAGATATGCGAGA 1741  
Qy 557 GATCCCGGAAATCACTCAAGAACCTTTGGAACATCAACCAACATACACACTAC 616  
Db 1742 GATCCCGGAAATCACTCAAGAACCTTTGGAACATCAACCAACATACACACTAC 1801  
Qy 617 ACTCAAGATCAGATGATCTTATTTGGGCTTTGGAACAGGCTTG 663  
Db 1802 ACTCAAGATCAGATGATCTTATTTGGGCTTTGGAACAGGCTTG 1848

## RESULT 9

AA221228  
ID AA221228 standard; cDNA; 2429 BP.  
AC AA221228;

DT 22-NOV-1999 (first entry)

DE Human CG1CE long form cDNA sequence.

KW CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;  
age-related macular dystrophy; ss.

OS Homo sapiens.

PH Key Location/Qualifiers  
FT CDS 105..1412  
/\*tag= a

/product= "CG1CE short form protein"

W09943695-A1.

BD 02-SEP-1999.

PE 22-FEB-1999; 99WO-US003790.

PR 25-FEB-1998; 98US-0075941P.  
PR 18-DEC-1998; 98US-0112926P.

PA (MERI ) MERCK & CO INC.  
(OTUP-) UNIV UPPSALA.

PI Petrukhin K, Caskey CT, Metzger M, Wadelius C;  
XX WPI; 1999-540560/45.

DR P-PSDB; AAY29954.

PT Human and mouse polynucleotides encoding CG1CE polypeptides.  
XX Human 2; Fig 4; 67bp; English.

CC The present sequence represents the human CG1CE cDNA sequence, which when  
mutated is responsible for Best's macular dystrophy (BMD).  
CC Polynucleotides encoding CG1CE are useful for diagnosing whether a  
patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE  
proteins are useful for identifying activators and/or inhibitors of these

CC proteins, in order to treat BMD. The CG1CE gene offers a simpler and  
CC cheaper method of diagnosing BMD without the need for the presence of the  
CC patient. The gene may also be useful to discovering the genetic cause of  
CC age-related macular dystrophy

XX SQ Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;

Query Match 29.0%; Score 642.2; DB 2; Length 2429;  
Best Local Similarity 99.5%; Pred. No. 7.1e-180;

Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 17 CAGCCTGAACAAAGAGAGATGAGTTCAGGCCCAATCAAGAGAGAGAGAGATGCTCA 76  
Db 1405 CAGCCTGAACAAAGAGAGATGAGTTCAGGCCCAATCAAGAGAGAGAGATGCTCA 1464  
Qy 77 CGCTGGATATATGGCCGCTTCTTCAAGCCTGCAATGATCAATCTCCAGAGGC 136  
Db 1465 CGCTGGATATATGGCCGCTTCTTCAAGCCTGCAATGATCAATCTCCAGAGGC 1524  
Qy 137 AAACCTGAAGACCAACTACTGTCGAGGAGGAGATCCCTTCCAGAGGCTTCC 196  
Db 1525 AAACCTGAAGACCAACTACTGTCGAGGAGGAGATCCCTTCCAGAGGCTTCC 1584  
Qy 197 CAAAAACCAAGGACGAGCCAAACAGAGCTTGAAGGAGGAGAGAGAGAGAGAGAGAG 256  
Db 1585 CAAAAACCAAGGACGAGCCAAACAGAGCTTGAAGGAGGAGAGAGAGAGAGAGAGAG 1644  
Qy 257 GCTTAAGGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATGAGAGGCGAGGCTTCAAG 316  
Db 1645 GCTTAAGGCTGTGAGCGCTTCAAGTGTGCGCCCACTGATGAGAGGCGAGGCTTCAAG 1704  
Qy 317 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 376  
Db 1705 TGCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTAGAACCATCAGCGCC 1764  
Qy 377 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAAGCTTAAAGCTGTAG 436  
Db 1765 GTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAAGCTTAAAGCTGTAG 1824  
Qy 437 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGGATGGGCTTATGAGCA 496  
Db 1825 TTCTGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGGATGGGCTTATGAGCA 1884  
Qy 497 CCCAGAAGTATCTCAAGTGAAGGAAAGCTTGAAGCTTAACTGAGATATGCGAGA 556  
Db 1885 CCCAGAAGTATCTCAAGTGAAGGAAAGCTTGAAGCTTAACTGAGATATGCGAGA 1944  
Qy 557 GATCCCGGAAATCACTCAAGAACCTTTGGAACATCAACCAACATACACACTAC 616  
Db 1945 GATCCCGGAAATCACTCAAGAACCTTTGGAACATCAACCAACATACACACTAC 2004  
Qy 617 ACTCAAGATCAGATGATCTTATTTGGGCTTTGGAACAGGCTTG 663  
Db 2005 ACTCAAGATCAGATGATCTTATTTGGGCTTTGGAACAGGCTTG 2051

## RESULT 10

ADA44960  
ID ADA44960 standard; cDNA; 1238 BP.

AC ADA44960;

DT 20-NOV-2003 (first entry)

DE Human polynucleotide #2.

KW Human; gene; ss; genetic disorder; genetic fingerprinting;  
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
KW insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;  
KW periodontal disease; bone fracture; cartilage damage;  
KW central nervous system disorder; Alzheimer's disease;  
KW Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;  
KW carbohydrate source.

XX OS Homo sapiens .  
XX PN US2003044935-A1.  
XX PD 06-MAR-2003.  
XX PF 21-DEC-2000; 2000US-00746783.  
XX  
PR 11-JUN-1997; 97US-0086236P.  
PR 12-JUN-1997; 97US-0086234P.  
PR 08-JUL-1997; 97US-0092115P.  
PR 08-SEP-1997; 97US-0093045P.  
PR 02-OCT-1997; 97US-0090100P.  
PR 27-OCT-1997; 97US-00958304.  
PR 07-NOV-1997; 97US-0090111P.  
PR 05-JUN-1998; 98US-0009272Z.  
PR 11-JUN-1998; 98US-00096287.  
PR 17-JUN-1998; 98US-00098588.  
PR 04-AUG-1998; 98US-00130189.  
PR 08-SEP-1998; 98US-00149633.  
PR 01-OCT-1998; 98US-00165860.  
PR 04-NOV-1998; 98US-00185936.  
XX  
PA (JACO/) JACOBS K.  
PA (MCCO/) MCCOY J M.  
PA (LVAL/) LA VALLIE E R.  
PA (COLL/) COLLINS-RACIE L A.  
PA (EVAN/) EVANS C.  
PA (MERB/) MERBERG D.  
PA (TREAC/) TREACY M.  
PA (SPAUL/) SPAULDING V.  
XX  
PI Jacobs K, Mccoy JM, La Vallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Spaulding V;  
XX  
DR MPI; 2003-521754/49.  
DR P-PSDB; ADA44961.  
XX  
PT New polypeptides and polynucleotides having biological activities, useful  
PT as nutritional sources or supplements, or for treating e.g. autoimmune  
PT diseases, cancers, bone fractures or damages, or central nervous system  
PT disorders.  
XX  
PS Claim 12; Page 109; 288bp; English.  
XX  
CC The invention relates to human polynucleotides and the polypeptides they  
CC encode. The polynucleotides can be used to express recombinant proteins  
CC for analysis, characterisation or therapeutic use, as markers for tissues  
CC in which the corresponding protein is expressed, as molecular weight  
CC markers on Southern gels, as chromosome markers or tags to identify  
CC chromosomes or to map related gene positions, to compare with endogenous  
CC DNA sequences in patients to identify potential genetic disorders, as  
CC probes to hybridise and discover novel related DNA sequences, as a source  
CC of information to derive PCR primers for genetic fingerprinting, to raise  
CC anti-protein antibodies and in gene therapy. The proteins can be used to  
CC raise antibodies or to elicit another immune response, as reagents in  
CC assays designed to quantitatively determine levels of the protein in  
CC biological fluids, as markers for tissues in which the corresponding  
CC protein is preferentially expressed and to treat autoimmune disorders  
CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent  
CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal  
CC diseases, bone fractures, cartilage damage, central nervous system  
CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.  
CC The proteins and polynucleotides are also useful as nutritional sources  
CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This  
CC sequence represents a human polynucleotide of the invention.  
XX  
SQ Sequence 1238 BP; 387 A; 336 C; 266 G; 249 T; 0 U; 0 Other;

QY 17 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 76  
|||  
Db 223 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 281  
|||  
QY 77 CGCTGGCATCATTTGGCCGCTTCTCTAGAGCCTGCAAGTCCCATGATCACCATCTCCAGAGC 136  
|||  
Db 282 CGCTGGCATCATTTGG-CGCTTCTTAGGCTGCAAGTCCCATGATCACCATCTCCAGAGC 340  
|||  
QY 137 AAACCTGAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCTCAGAGAGCCTGCC 196  
|||  
Db 341 AAACCTGAAGACC-AACTACTGTGGCCCAAGAGGAATCCCTTCTCAGAGAGCCTGCC 399  
|||  
QY 197 CAAAAACCAAGGACGACCAAGACGTTAGGGGCGAAGAGACAAAGGCTTGAA 256  
|||  
Db 400 CAAAAACCAAGGAC-GCCAAACGAACGTTAGGGGCGAAGAGACAAAGGCTTGAA 458  
|||  
QY 257 GCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGGCTACTAG 316  
|||  
Db 459 GCTTAAAGCTGTGGA-GCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGGCTACTAG 517  
|||  
QY 317 TGCCCCACAGAGCGCCCTCAGCCCACTCCCATGTTCTTCCCTTGAACCATCAGGCC 376  
|||  
Db 518 TGCCCCACAGAGC-CCCTCAGCCCACTCCCATGTTCTTCCCTTGAACCATCAGGCC 576  
|||  
QY 377 GTCAAGCTTCAAGTGTCAAGGACATGACCAAAAGACCTTAAAGCTGTAG 436  
|||  
Db 577 GTCAAGCTTCAAG-GTCAAGGACATGACCAAAAGACCTTAAAGCTGTAG 635  
|||  
QY 437 TTCTGGGGCCAAAGAAAGTTTGAATTGCTCTTCAGAGAGGAGTGGGCTTGAATGAGCA 496  
|||  
Db 636 TTCTGGGGCCAAAG-AAAGTTTGAATTGCTCTTCAGAGAGGAGTGGGCTTGAATGAGCA 694  
|||  
QY 497 CCCAGAAGTATCTCAAGTGAAGGAAACTGTGAGTTTAACCTGACGATATGCCAGA 556  
|||  
Db 695 CCCAGAAGTATCTC-AGTGAAGGAAACTGTGAGTTTAACCTGACGATATGCCAGA 753  
|||  
QY 557 GATCCCCGAAATCACCTCAAGAACCTTTGAAACATCCAAACCAATACACACTAC 616  
|||  
Db 754 GATCCCCGAAATCA-CTCAAGAACCTTTGAAACATCCAAACCAATACACACTAC 812  
|||  
QY 617 ACTCAAGATCACATGATCTTATTTGGGCTTGAAGAAACAGGCTG 663  
|||  
Db 813 ACTCAAGATCACAT-GATCTTATTTGGGCTTGAAGAAACAGGAGTG 858  
|||  
RESULT 11  
ID ABA14556 standard; DNA; 7108 BP.  
XX  
AC ABA14556;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Human nervous system related polynucleotide SEQ ID NO 6887.  
XX  
KW Human; nootropic; neuroprotective; cyrostatic; dermatological; viroicide;  
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulniterary;  
KW antiparkinsonian; antistroke; antianemic; antiairthritis; cancer;  
KW antineumatic; hepatocytic; cerebroprotective; antiinflammatory;  
KW antidiabetic; antidiabetic; antidiabetic; antidiabetic; antifungal;  
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.  
XX  
OS Homo sapiens.  
XX  
PN MO200159063-A2.  
XX  
PD 16-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001334.  
XX  
PX 31-JAN-2000; 2000US-0179065P.

PR	04-FEB-2000	2000US-0180678P
PR	24-FEB-2000	2000US-0184664P
PR	02-MAR-2000	2000US-0196350P
PR	16-MAR-2000	2000US-0198974P
PR	17-MAR-2000	2000US-0190076P
PR	18-MAR-2000	2000US-0198123P
PR	19-MAR-2000	2000US-0205515P
PR	07-JUN-2000	2000US-0209467P
PR	28-JUN-2000	2000US-0214866P
PR	30-JUN-2000	2000US-0215135P
PR	07-JUL-2000	2000US-0216647P
PR	11-JUL-2000	2000US-0217487P
PR	14-AUG-2000	2000US-0225513P
PR	14-AUG-2000	2000US-0225514P
PR	14-AUG-2000	2000US-0225666P
PR	14-AUG-2000	2000US-0225677P
PR	14-AUG-2000	2000US-0225688P
PR	14-AUG-2000	2000US-0225747P
PR	14-AUG-2000	2000US-0224511P
PR	14-AUG-2000	2000US-0225515P
PR	14-AUG-2000	2000US-0225757P
PR	14-AUG-2000	2000US-0225758P
PR	14-AUG-2000	2000US-0225759P
PR	22-AUG-2000	2000US-0226679P
PR	22-AUG-2000	2000US-0226681P
PR	22-AUG-2000	2000US-0226686P
PR	23-AUG-2000	2000US-0227189P
PR	30-AUG-2000	2000US-0227099P
PR	01-SEP-2000	2000US-0228924P
PR	01-SEP-2000	2000US-0229287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229344P
PR	05-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	06-SEP-2000	2000US-0228513P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231243P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231443P
PR	08-SEP-2000	2000US-0231444P
PR	08-SEP-2000	2000US-0232080P
PR	12-SEP-2000	2000US-0232081P
PR	14-SEP-2000	2000US-0231968P
PR	14-SEP-2000	2000US-0232337P
PR	14-SEP-2000	2000US-0232338P
PR	14-SEP-2000	2000US-0232339P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0232403P
PR	14-SEP-2000	2000US-0233063P
PR	21-SEP-2000	2000US-0233065P
PR	21-SEP-2000	2000US-0234223P
PR	25-SEP-2000	2000US-0234274P
PR	25-SEP-2000	2000US-0234977P
PR	26-SEP-2000	2000US-0234988P
PR	27-SEP-2000	2000US-0235534P
PR	27-SEP-2000	2000US-0235535P
PR	29-SEP-2000	2000US-0236127P
PR	29-SEP-2000	2000US-0236127P
PR	29-SEP-2000	2000US-0236168P
PR	29-SEP-2000	2000US-0236169P
PR	02-OCT-2000	2000US-0236602P
PR	02-OCT-2000	2000US-0237037P

PR	02-OCT-2000	2000US-0237038P.
PR	02-OCT-2000	2000US-0237039P.
PR	02-OCT-2000	2000US-0237040P.
PR	13-OCT-2000	2000US-0239935P.
PR	13-OCT-2000	2000US-0239937P.
PR	20-OCT-2000	2000US-0240960P.
PR	20-OCT-2000	2000US-0241785P.
PR	20-OCT-2000	2000US-0241786P.
PR	20-OCT-2000	2000US-0241787P.
PR	20-OCT-2000	2000US-0241808P.
PR	20-OCT-2000	2000US-0241809P.
PR	20-OCT-2000	2000US-0242221P.
PR	01-NOV-2000	2000US-0244617P.
PR	08-NOV-2000	2000US-0246474P.
PR	08-NOV-2000	2000US-0246475P.
PR	08-NOV-2000	2000US-0246476P.
PR	08-NOV-2000	2000US-0246477P.
PR	08-NOV-2000	2000US-0246478P.
PR	08-NOV-2000	2000US-0246523P.
PR	08-NOV-2000	2000US-0246524P.
PR	08-NOV-2000	2000US-0246525P.
PR	08-NOV-2000	2000US-0246527P.
PR	08-NOV-2000	2000US-0246528P.
PR	08-NOV-2000	2000US-0246532P.
PR	08-NOV-2000	2000US-0246609P.
PR	08-NOV-2000	2000US-0246610P.
PR	08-NOV-2000	2000US-0246611P.
PR	17-NOV-2000	2000US-0246613P.
PR	17-NOV-2000	2000US-0249207P.
PR	17-NOV-2000	2000US-0249208P.
PR	17-NOV-2000	2000US-0249209P.
PR	17-NOV-2000	2000US-0249210P.
PR	17-NOV-2000	2000US-0249211P.
PR	17-NOV-2000	2000US-0249212P.
PR	17-NOV-2000	2000US-0249213P.
PR	17-NOV-2000	2000US-0249214P.
PR	17-NOV-2000	2000US-0249215P.
PR	17-NOV-2000	2000US-0249216P.
PR	17-NOV-2000	2000US-0249217P.
PR	17-NOV-2000	2000US-0249218P.
PR	17-NOV-2000	2000US-0249244P.
PR	17-NOV-2000	2000US-0249245P.
PR	17-NOV-2000	2000US-0249264P.
PR	17-NOV-2000	2000US-0249265P.
PR	17-NOV-2000	2000US-0249297P.
PR	17-NOV-2000	2000US-0249298P.
PR	17-NOV-2000	2000US-0249300P.
PR	01-DEC-2000	2000US-0250391P.
PR	01-DEC-2000	2000US-0251160P.
PR	05-DEC-2000	2000US-0251030P.
PR	05-DEC-2000	2000US-0251988P.
PR	06-DEC-2000	2000US-0256719P.
PR	08-DEC-2000	2000US-0251856P.
PR	08-DEC-2000	2000US-0251866P.
PR	08-DEC-2000	2000US-0251869P.
PR	08-DEC-2000	2000US-0251989P.
PR	11-DEC-2000	2000US-0254097P.
PR	05-JAN-2001	2001US-0259678P.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM,	
XX		
DR	WPI; 2001-541565/60.	
PT	Nucleic acids encoding 3224 human nervous system antigen polypeptides,	
XX	useful for preventing, diagnosing and/or treating nervous system cancers	
XX	and metastases.	





QY 1633 TCGTGTGTCCTGTTAGTCCCAACGAGAGGTTGAGGAGAAATTGCTTGACCCAGGA 1692  
 Db 321 TCGTGTGTCCTGTTAGTCCCAACGAGAGGTTGAGGAGAAATTGCTTGACCCAGGA 380  
 QY 1693 GGTGTGTGTTGAGTGAATGAGCAATGCAATCCAGCTGGGCGACGAGTGAAGACTG 1752  
 Db 381 GGTGTGTGTTGAGTGAATGAGCAATGCAATCCAGCTGGGCGACGAGTGAAGACTG 440  
 QY 1753 TCTCAAAAAAAAAAAAAA 1771  
 Db 441 TCTCAAAAAAAAAAAAAA 459

RESULT 13  
 ID AAV86976 standard; cDNA; 424 BP.  
 XX AAV86976;  
 AC AAV86976;  
 XX 27-APR-1999 (first entry)  
 DT 27-APR-1999 (first entry)  
 DE EST clone BK112.  
 XX  
 KW Expressed sequence tag; secreted protein; haematopoiesis regulator;  
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;  
 KW chemokinesis; chemokinesis; haemostasis; gene therapy; thrombolysis;  
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO845435-A2.  
 XX  
 PD 15-OCT-1998.  
 XX  
 PF 10-APR-1998; 98WO-US006954.  
 XX  
 PR 10-APR-1997; 97US-00835913.  
 XX  
 PA (GENE) GENETICS INST INC.  
 XX  
 PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M,  
 PI Spaulding V, Agostino MJ;  
 XX  
 DR WPI; 1999-070076/06.  
 XX  
 XX New polynucleotides encoding human secreted proteins - derived from e.g.  
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,  
 PT pituitary, retina and colon cDNA libraries.  
 XX  
 PS Claim 1; Page 411; 633pp; English.  
 XX  
 CC This sequence represents an expressed sequence tag (EST), and is a  
 CC polynucleotide of the invention. The polynucleotides of the invention are  
 CC all secreted EST sequences isolated from a variety of human tissue  
 CC sources. The EST sequences and proteins encoded by them are predicted to  
 CC have useful biological activities and proteins which would make them suitable for  
 CC treating, preventing or ameliorating medical conditions in humans and  
 CC animals, although no supporting data is given. Suggested activities  
 CC include nutritional activity, immune stimulating or suppressing activity,  
 CC haematopoiesis regulating activity, tissue growth activity,  
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC activity, thrombolytic activity, receptor/ligand activity, anti-inflammatory  
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition  
 CC activity. The EST sequences are also stated to be useful for gene therapy

Sequence 424 BP; 110 A; 124 C; 95 G; 95 T; 0 U; 0 Other;  
 Query Match 18.0%; Score 399.4; DB 2; Length 424;  
 Best Local Similarity 99.8%; Pred. No. 5e-108;  
 Matches 400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 426 AAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCC 485

Db 21 AAGACTGTGAGTTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGGATGGGCC 80  
 QY 486 TTGATGAGGACACCCAGAAAGTATCTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACG 545  
 Db 81 TTGATGAGGACACCCAGAAAGTATCTCAAGTGAAGAGAGAACTGTGAGTTTAACTGACG 140  
 QY 546 GATATGCCAGAGATCCCGGAAAAATCACTCAAGAAACCTTTGGAAACATACCAACCAAC 605  
 Db 141 GATATGCCAGAGATCCCGGAAAAATCACTCAAGAAACCTTTGGAAACATACCAACCAAC 200  
 QY 606 ATACACACTACACTCAAGATGCAATGATGATCTTATGAGGCTTTGAAAAACAGTCTGTC 665  
 Db 201 ATACACACTACACTCAAGATGCAATGATGATCTTATGAGGCTTTGAAAAACAGTCTGTC 260  
 QY 666 CTCACCTGTAACGAGGAGGACATGATGCTGCTGTGCCCCACCCAGCTTCCCTGCTTG 725  
 Db 261 CTCACCTGTAACGAGGAGGACATGATGCTGCTGTGCCCCACCCAGCTTCCCTGCTTG 320  
 QY 726 AGCCTACCTTCTCTCCCAATTTCTTAAAGGTTCCATCACTGCGACAGACACTGACCTA 785  
 Db 321 AGCCTACCTTCTCTCCCAATTTCTTAAAGGTTCCATCACTGCGACAGACACTGACCTA 380  
 QY 786 CGCCGACGACTGAGCTTGGGGTATATATCTTGGGCACTTGCAC 826  
 Db 381 CGCCGACGACTGAGCTTGGGGTATATATCTTGGGCACTTGCAC 421

RESULT 14  
 ID ACA03889/C  
 XX ACA03889 standard; cDNA; 1326 BP.  
 XX  
 AC ACA03889;  
 XX  
 DT 27-MAY-2003 (first entry)  
 XX  
 DE cDNA upregulated in senescent cells Incyte ID NO: 1138151.2.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2002192678-A1.  
 XX  
 PD 19-DEC-2002.  
 XX  
 PF 07-FEB-2002; 2002US-00071766.  
 XX  
 PR 09-FEB-2001; 2001US-0268380P.  
 XX  
 PA (CHEN/) CHEN H.  
 XX  
 PI Chen H;  
 XX  
 DR WPI; 2003-328858/31.  
 XX

PT New combination comprising cDNAs or their complements, useful for  
 PT detecting changes in expression of genes encoding proteins associated  
 PT with senescence, and in diagnosing, staging or treating proliferative  
 PT diseases, e.g. cancer.  
 XX  
 PS Disclosure; Page 27; 195pp; English.  
 XX  
 CC The invention relates to a combination comprising a plurality of cDNAs,



CC or their complements that are differentially expressed in cancer and  
CC other proliferative disorders. The combination is useful in detecting  
CC changes in expression of genes encoding proteins that are associated with  
CC senescence and in diagnosing, staging, treating, or monitoring the  
CC progression or treatment of subjects with proliferative diseases such as  
CC cancer (e.g. adenocarcinoma, leukemia, lymphoma, melanoma, myeloma,  
CC sarcoma, leiomyosarcoma, cancer of the adrenal gland, bladder, bone,  
CC bone marrow, brain, breast, cervix, colon, oesophagus, gall bladder,  
CC ganglia, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid  
CC gland, penis, prostate, salivary glands, skin, small intestine, spleen,  
CC stomach, testis, thymus, thyroid and uterus. The present invention  
CC represents cDNA of genes that are upregulated in senescent cells

XX Sequence 1326 BP; 338 A; 328 C; 315 G; 345 T; 0 U; 0 Other;

Query Match 15.3%; Score 339.4; DB 7; Length 1326;  
Best Local Similarity 99.7%; Pred. No. 7.2e-90;  
Matches 340; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1876 AGCAGATTCCTAACCTGCTTCTTATATGGGGATGCTTGCCAGCCAGGTCCTCACTGTGT 1935  
DB 1326 AGCAGATTCCTAACCTGCTTCTTATATGGGGATGCTTGCCAGCCAGGTCCTCACTGTGT 1267

QY 1936 GTAACACGAGAGAGACACTGATCCAGTCACAGCCATAGCTGTCCACATGAAAGATG 1995  
DB 1266 GTAACACGAGAGAGACACTGATCCAGTCACAGCCATAGCTGTCCACATGAAAGATG 1207

QY 1996 TCCTCAACAAGCCTGAATCAATATGCTTAATATGATTAATAATCCCAAGATCTTCAAG 2055  
DB 1206 TCCTCAACAAGCCTGAATCAATATGCTTAATATGATTAATAATCCCAAGATCTTCAAG 1147

QY 2056 CCTTATATGCTTTTATTTATTAATAAAGTGAAGCTGAGCACTGGAACATTTGAAACATTT 2115  
DB 1146 CCTTATATGCTTTTATTTATTTAATAAAGTGAAGCTGAGCACTGGAACATTTGAAACATTT 1087

QY 2116 AACTCAGACTCTGATTCAGAGTCGGGAAACCTTAGTTCTATCTGAATCCCAAGACAGCA 2175  
DB 1086 AACTCAGACTCTGATTCAGAGTCGGGAAACCTTAGTTCTATCTGAATCCCAAGACAGCA 1027

QY 2176 CACCTTAGTATCTGCTGCCAAACTATATAGTTTAATAATNC 2216  
DB 1026 CACCTTAGTATCTGCTGCCAAACTATATAGTTTAATAATNC 986

RESULT 15  
AAC20667  
ID AAC20667 standard; cDNA; 356 BP.  
XX  
AC AAC20667;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 24742.  
XX  
KM Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
OS  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-00200610.  
XX  
PR 26-FEB-1999; 99US-0122487P.  
XX  
PA (GBST ) GENSET.  
XX  
PI Dumas Mline Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX  
XX Claim 1; SEQ ID NO 24742; 71bp + Sequence Listing; English.  
XX

XX The present sequence is one of a large number of 5' ESTs derived from  
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively  
XX identified within the present sequence. The 5' ESTs were prepared from  
XX total human RNA or polyA+ RNAs derived from 30 different tissues. EST  
XX sequences usually correspond mainly to the 3' untranslated region (UTR)  
XX of the mRNA because they are often obtained from oligo-dT primed cDNA  
XX libraries. Such ESTs are not well suited for isolating cDNA sequences  
XX derived from the 5' ends of mRNAs and even in those cases where longer  
XX cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'  
XX ESTs are derived from mRNAs with intact 5' ends and can therefore be used  
XX to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in  
XX diagnostic, forensic, gene therapy and chromosome mapping procedures.  
XX They are used to obtain upstream regulatory sequences and to design  
XX expression and secretion vectors

XX Sequence 356 BP; 103 A; 69 C; 90 G; 88 T; 0 U; 6 Other;

Query Match 14.4%; Score 318; DB 3; Length 356;  
Best Local Similarity 95.8%; Pred. No. 7.8e-84;  
Matches 342; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1331 CACGTGGAGAGAAAGTAACTTCCCTTTCTGATTTCTCAAGACATTTCTTACGGGTC 1390  
DB 1 CACGTGGAGAGAAAGTAACTTCCCTTTCTGATTTCTCAAGACATTTCTTACGGGTC 59

QY 1391 AGAACACGACGATTTATGATTTGAACCTTAAAGGCAACAATTCANTCTGCTCTA 1450  
DB 60 AGAACACGACGATTTATGATTTGAACCTTAAAGGCAACAATTCANTCTGCTCTA 119

QY 1451 GGCTAAGACAGAACTTGGCAACATCTGTGCTGCTTTCAGCAAGAGATTTATATTTA 1510  
DB 120 GGGTGAACAGAACTTGGCAACATCTGTGCTGCTTTCAGCAAGAGATTTATATTTA 179

QY 1511 AGAATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATACACAGAGAGTCAAGATTGAG 1570  
DB 180 AGAATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATACACAGAGAGTCAAGATTGAG 239

QY 1571 ACCAAGCTGGCAACATGATGAAACCCATCTTACCAAAAAAATAC--AAATCAGCTG 1628  
DB 240 ACCAAGCTGGCAACATGATGAAACCCATCTTACCAAAAAAATACAAATTCAGCTG 299

QY 1629 GCCGTGCTGGTGTGCTGTAGTCCCAACGAGAGGTTGAGGAGAAATTTGCTTGA 1685  
DB 300 GCCGTGCTGGTGTGCTGTAGTCCCAACGAGAGGTTGAGGAGAAATTTGCTTGA 356

Search completed: March 25, 2004, 03:23:35  
Job time : 861 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 03:09:19 ; Search time 162 Seconds

(without alignments)  
7591.180 Million cell updates/sec

Title: US-09-622-964-1\_COPY\_13900\_16115

Perfect score: 2216  
Sequence: 1 actctcgtctcttcacgc.....craatgagtttaataaac 2216

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/ptodaca/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/ptodaca/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/ptodaca/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/ptodaca/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/ptodaca/2/ina/PTUS COMB.seq: \*  
6: /cgn2\_6/ptodaca/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133.2	6.0	65042	4 US-09-784-316-3	Sequence 3, Appl1
2	123	5.6	7720	3 US-09-318-448-5	Sequence 5, Appl1
3	121.6	5.5	36159	4 US-09-749-588-3	Sequence 3, Appl1
4	119.4	5.4	49312	4 US-09-671-317-485	Sequence 485, App
5	117.2	5.3	2867	4 US-09-402-532-38	Sequence 38, Appl1
6	117.2	5.3	21234	4 US-09-810-671-3	Sequence 3, Appl1
7	117.2	5.3	21234	4 US-10-109-854-3	Sequence 42, Appl1
8	117	5.3	44848	4 US-09-435-739-42	Sequence 1, Appl1
9	117	5.3	162450	4 US-09-345-882-1	Sequence 1, Appl1
10	116.8	5.3	90541	4 US-09-759-359A-3	Sequence 1, Appl1
11	116.8	5.3	92139	4 US-09-918-686-1	Sequence 1, Appl1
12	116.2	5.2	62804	4 US-09-800-960-3	Sequence 3, Appl1
13	116.2	5.2	62804	4 US-10-096-960-3	Sequence 3, Appl1
14	115.8	5.2	18853	4 US-09-820-005-3	Sequence 3, Appl1
15	115.4	5.2	2296	4 US-09-851-896-11	Sequence 11, Appl1
16	115.2	5.2	99500	4 US-09-798-096-10	Sequence 10, Appl1
17	115.2	5.2	168575	4 US-09-426-290-1	Sequence 1, Appl1
18	114	5.1	246240	2 US-08-724-394A-20	Sequence 20, Appl1
19	114	5.1	246240	2 US-08-724-394A-21	Sequence 21, Appl1
20	114	5.1	246240	2 US-08-724-394A-22	Sequence 22, Appl1
21	113.8	5.1	2236	3 US-08-829-525-23	Sequence 23, Appl1
22	113.8	5.1	2236	3 US-08-609-583A-23	Sequence 23, Appl1
23	113.8	5.1	2236	3 US-08-937-399-23	Sequence 23, Appl1
24	113.8	5.1	2236	4 US-09-310-367-23	Sequence 23, Appl1
25	113.8	5.1	2236	4 US-09-032-337-23	Sequence 23, Appl1
26	113.8	5.1	2236	4 US-09-464-231-23	Sequence 23, Appl1
27	113.8	5.1	53332	4 US-09-801-861-3	Sequence 3, Appl1

C 28	113.6	5.1	72604	4 US-09-268-992-7	Sequence 7, Appl1
C 29	113.6	5.1	72604	4 US-09-657-474-7	Sequence 7, Appl1
C 30	113.6	5.1	87350	3 US-08-781-891-79	Sequence 79, Appl1
C 31	113.6	5.1	87350	4 US-09-618-166-79	Sequence 79, Appl1
C 32	113.6	5.1	87353	4 US-09-791-211-3	Sequence 3, Appl1
C 33	113	5.1	18596	3 US-09-318-448-11	Sequence 11, Appl1
C 34	113	5.1	21234	4 US-09-810-671-3	Sequence 3, Appl1
C 35	113	5.1	21234	4 US-10-109-854-3	Sequence 3, Appl1
C 36	112.8	5.1	21721	4 US-09-269-939A-41	Sequence 41, Appl1
C 37	112.8	5.1	22976	4 US-09-269-939A-19	Sequence 19, Appl1
C 38	112.8	5.1	23187	4 US-09-499-522-1	Sequence 1, Appl1
C 39	112.8	5.1	64467	4 US-09-803-671B-3	Sequence 3, Appl1
C 40	112.8	5.1	70000	4 US-09-851-896-3	Sequence 3, Appl1
C 41	112.8	5.1	162450	4 US-09-345-882-1	Sequence 1, Appl1
C 42	112.6	5.1	2932	4 US-09-016-434-1419	Sequence 1419, Ap
C 43	112	5.1	6235	3 US-09-305-384-5	Sequence 5, Appl1
C 44	112	5.1	6235	4 US-09-525-160B-6	Sequence 6, Appl1
C 45	112	5.1	6679	3 US-09-305-384-1	Sequence 1, Appl1

## ALIGNMENTS

RESULT 1  
US-09-784-316-3/C  
; Sequence 3, Application US/09784316  
; Patent No. 6461843  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming-Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; FILE REFERENCE: C1001139  
; CURRENT APPLICATION NUMBER: US/09/784,316  
; CURRENT FILING DATE: 2001-02-16  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 65042  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(65042)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-784-316-3

Query Match 6.0%; Score 133.2; DB 4; Length 65042;  
Best Local Similarity 70.6%; Pred. No. 5.1e-31;  
Matches 209; Conservative 0; Mismatches 78; Indels 9; Gaps 2;

QY	1527	CTGGGTGTGAGGCAAGTGAATCAGAGAGTTCAGAGATTGAGACCAACTGGCCACACA	1586
DB	35972	CGGAGAGCGGAGGCGGCGGATCACTGAGTTCAGAGATCCAGACCAAGCTGGCCACACA	35913
QY	1587	TGATGAAGACCCATCTCTCAAAAAAATACAAATACAGTGGCCGTGTGTCCTG	1646
DB	35912	TGATGAAGACCCATCTCTCAAAAAAATACAAATACAGTGGCCGTGTGTCCTG	35853
QY	1647	TGATCCCAAC---GCAGAGGTGAGGAGAGATTCCTTGAACCCAGAGAGTGTGCTT	1702
DB	35852	TGATCCCAAC---GCAGAGGTGAGGAGAGATTCCTTGAACCCAGAGAGTGTGCTT	35793
QY	1703	GCAG---TGAGATTGAGCAACTGCAATTCACGCTGGGCGACGAGTGAAGTGTCTCA	1757
DB	35792	GCAGTGAAGTGAAGTGTGCTCACTGCACTGAGCTGGGCGACGAGTGAAGTGTCTCA	35733
QY	1758	AAAAAAAAAAAAAAAAAGATGCTCTCAACTTGGCCCTCTACAGCAACATTTGGT	1813
DB	35732	CAAAAAAAAAAAAAAAAAAGATGCTCTCAACTTGGCCCTCTACAGCAACATTTGGT	35677

RESULT 2

QY 1536 GAGGCAAGTAATCA CAGAGGCGAGAGTTTGACCACTGGCCAA CATGATGAAC 1595  
Db 3007 GAGTGGGTGATCATCTGAGTCCAGAACTTTGAGACCAAGCTGGCCAACTAGGAAC 2948  
QY 1596 CCCATCTCACCAAAAAAATACAATCACTGCGCCGTGTGTGTGCTGTGATGCCA- 1654  
Db 2947 CCTGTCTCTACAAAAATACAAAAAGAGCTGGGTGCGGTGTGTGTGCTGTAGTCCAG 2888  
QY 1655 --ACCCAGAGGTTAAG- GGAGATTTGCTTGAACCCAGAGGAGTGGTGACAG- 1708  
Db 2887 CTACGTGGAGGCTGAGCAGAGAGATCACTTGATCCAGAGGTGAGGTTACAGTGA 2828  
QY 1709 ---AGATTGACCACTGCAATCAAGCCCTGGCGCAGACGAGTGAGACTGTCTCAAAAAA 1765  
Db 2827 CCAAGATTGTCCACTTGCACTCTGAGCCTGGGTGACAGAGGAGACTCCATCTCAAAAAA 2768  
QY 1766 AAAAAAGATGATCTCAACTCTTGGCCCTCTACTGCAACATTTTGGTATTGG 1818  
Db 2767 AAAAAAAGTTTATAAAAAATGTTTAACTGTGGGGCGCTGACGTGTGGTTTTG 2715

```

RESULT 3
US-09-749-588-3/c
; Sequence 3, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS,
; TITLE OF INVENTION: THREEOF
; FILE REFERENCE: C1001068
; CURRENT APPRECIATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(36159)
; OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

```

Query Match	5.5%	Score 121.6;	DB 4;	Length 36159;
Best Local Similarity	78.1%	Pred. No. 1.9e-27;		
Matches 200; Conservative	0;	Mismatches 44;	Indels 12;	Gaps 4;
QY	1529	GGGTGTGAGGCAACTGAATCAGAGAGGTTCAGATTGAGACCAACTGGCCCAATG	1588	

Accession	Sequence	Position
Db	11721 GGAGCCGAGAGGCTGATCACCCTGAAGTCACAGAGTTTGAGACCGCCTGGCAACATG	11672
OY	1589 ATGAACCCCATCTCTACCAAAAAAATACAAATCACTGACCGCTGGTGTGTG--CTTG	1646
Db	11671 GTGAACCCCATCTCTCTCTAAATATTAATAAATACGACAGTGTGTGTGGCGGCACTG	11612
OY	1647 TAGTCCCA---ACGCAAGAGGTTGAGG--GGAGAAATGGCTTGAACCCAGAGGTGTGT	1701
Db	11611 TAATCCCAAGTACTCAGAGGCTGAGGCGAGGAGAAATGGCTTGAACCCAGAGAACGAGGT	11552
OY	1702 TGCAGTGAG----ATTGACAACTGCATTCAGGCTGGGGCGACCGAGTGAGACTGTCTC	1756
Db	11551 TGCAGTGAGCTGACACAGTGCCACTGTCACTTCAGGCTGGGTGACAGAGTAAAGACTCTGTC	11492
OY	1757 AAAAAAAAAAAAAAAAAA 1772	
Db	11491 TCAAAAAAAAAAAAAA 11476	

```

Sequence 485 Application US/09671317
Patent No. 6528260
GENERAL INFORMATION:
APPLICANT: Blumenfeld, Marta
APPLICANT: Chumakov, Ilya
APPLICANT: Bougueleret, Lydie
APPLICANT: Cohen, Annick
TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
FILE REFERENCE: 62, US3, CIP
CURRENT APPLICATION NUMBER: US/09/671,317
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US 09/536,178
PRIOR FILING DATE: 2000-03-23
PRIOR APPLICATION NUMBER: PCT/IB00/00403
PRIOR FILING DATE: 2000-03-24
PRIOR APPLICATION NUMBER: US 60/126,269
PRIOR FILING DATE: 1999-03-25
PRIOR APPLICATION NUMBER: US 60/131,961
PRIOR FILING DATE: 1999-04-30
NUMBER OF SEQ ID NOS: 977
SOFTWARE: Patent.pm
SEQ ID NO 485

```

```

1 ORGANISM: Homo sapiens
2 FEATURE:
3 NAME/KEY: misc feature
4 LOCATION: 5466..7466
5 OTHER INFORMATION: 5'regulatory region
6 NAME/KEY: exon
7 LOCATION: 7467..7725
8 OTHER INFORMATION: exon 1
9 NAME/KEY: exon
10 LOCATION: 20256..20355
11 OTHER INFORMATION: exon 2
12 NAME/KEY: exon
13 LOCATION: 36905..36975
14 OTHER INFORMATION: exon 3
15 NAME/KEY: exon
16 LOCATION: 45167..45248
17 OTHER INFORMATION: exon 4
18 NAME/KEY: exon
19 LOCATION: 45728..45965
20 OTHER INFORMATION: exon 5
21 NAME/KEY: misc feature
22 LOCATION: 45966..49312
23 OTHER INFORMATION: 3'regulatory region
24 NAME/KEY: allele
25 LOCATION: 7564
26 OTHER INFORMATION: 10-286-289 : polymorphic base G or C
27 NAME/KEY: allele

```

```
LOCATION: 7619
OTHER INFORMATION: 10-286-345 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 7649
OTHER INFORMATION: 10-286-375 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 17258
OTHER INFORMATION: 12-425-57 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 21590
OTHER INFORMATION: 12-421-135 : insertion of T
NAME/KEY: allele
LOCATION: 21595
OTHER INFORMATION: 12-421-140 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 36971
OTHER INFORMATION: 10-523-232 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45214
OTHER INFORMATION: 10-289-201 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45741
OTHER INFORMATION: 10-290-37 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 46029
OTHER INFORMATION: 10-290-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 46032
OTHER INFORMATION: 10-290-328 : deletion of G
NAME/KEY: primer bind
LOCATION: 7276..7294
OTHER INFORMATION: 10-286.pu
NAME/KEY: primer bind
LOCATION: 7676..7694
OTHER INFORMATION: 10-286.rp complement
NAME/KEY: primer bind
LOCATION: 16839..16856
OTHER INFORMATION: 12-425.rp
NAME/KEY: primer bind
LOCATION: 17297..17314
OTHER INFORMATION: 12-425.pu complement
NAME/KEY: primer bind
LOCATION: 21456..21474
OTHER INFORMATION: 12-421.pu
NAME/KEY: primer bind
LOCATION: 21886..21906
OTHER INFORMATION: 12-421.rp complement
NAME/KEY: primer bind
LOCATION: 36740..36758
OTHER INFORMATION: 10-523.pu
NAME/KEY: primer bind
LOCATION: 36997..37015
OTHER INFORMATION: 10-523.rp complement
NAME/KEY: primer bind
LOCATION: 45020..45037
OTHER INFORMATION: 10-289.pu
NAME/KEY: primer bind
LOCATION: 45413..45432
OTHER INFORMATION: 10-289.rp complement
NAME/KEY: primer bind
LOCATION: 45705..45724
OTHER INFORMATION: 10-290.pu
NAME/KEY: primer bind
LOCATION: 46104..46123
OTHER INFORMATION: 10-290.rp complement
NAME/KEY: primer bind
LOCATION: 7545..7563
OTHER INFORMATION: 10-286-289.mis
NAME/KEY: primer bind
LOCATION: 7565..7583
OTHER INFORMATION: 10-286-289.mis complement
NAME/KEY: primer bind
LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
NAME/KEY: primer bind
LOCATION: 7620..7638
OTHER INFORMATION: 10-286-345.mis complement
NAME/KEY: primer bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
NAME/KEY: primer bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
NAME/KEY: primer bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
NAME/KEY: primer bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
NAME/KEY: primer bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
NAME/KEY: primer bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
NAME/KEY: primer bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
NAME/KEY: primer bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
NAME/KEY: primer bind
LOCATION: 45195..45213
OTHER INFORMATION: 10-289-201.mis
NAME/KEY: primer bind
LOCATION: 45215..45233
OTHER INFORMATION: 10-289-201.mis complement
NAME/KEY: primer bind
LOCATION: 45722..45740
OTHER INFORMATION: 10-290-37.mis
NAME/KEY: primer bind
LOCATION: 45742..45760
OTHER INFORMATION: 10-290-37.mis complement
NAME/KEY: primer bind
LOCATION: 46010..46028
OTHER INFORMATION: 10-290-326.mis
NAME/KEY: primer bind
LOCATION: 46030..46048
OTHER INFORMATION: 10-290-326.mis complement
NAME/KEY: primer bind
LOCATION: 7552..7576
OTHER INFORMATION: 10-286-289.probe
NAME/KEY: misc binding
LOCATION: 7607..7631
OTHER INFORMATION: 10-286-345.probe
NAME/KEY: misc binding
LOCATION: 7637..7661
OTHER INFORMATION: 10-286-375.probe
NAME/KEY: misc binding
LOCATION: 17246..17270
OTHER INFORMATION: 12-425-57.probe
NAME/KEY: misc binding
LOCATION: 21583..21607
OTHER INFORMATION: 12-421-140.probe
NAME/KEY: misc binding
LOCATION: 36959..36983
OTHER INFORMATION: 10-523-232.probe
NAME/KEY: misc binding
LOCATION: 45202..45226
OTHER INFORMATION: 10-289-201.probe
NAME/KEY: misc binding
LOCATION: 45725..45753
OTHER INFORMATION: 10-290-37.probe
NAME/KEY: misc binding
LOCATION: 46017..46041
OTHER INFORMATION: 10-290-326.probe
```

US-09-671-317-485

Query Match  
Best Local Similarity 5.4%; Score 119.4; DB 4; Length 49312;  
Matches 205; Conservative 0; Mismatches 56; Indels 12; Gaps 4;

QY 1528 TGGGTGTGAGCAATGATCAAGAGCTGAGTTTGAACCACTTGCCCAAT 1587  
DB 16251 TGGAGGCTAAGGGGGGAGTACTTGTAGGTGAGGTTCAAGACCACTTGCCCAAT 16310  
QY 1588 GATGAACCCCATCTCTCAACAAAAATACAATAGCTGGCGGTGGG--TGTGCTT 1645  
DB 16311 GGTGAACCCCATCTCTCTCAACAAAAATACAATAGCTGGCGGTGGG--TGTGCTT 16370  
QY 1646 GTAGTCCCA--ACGAGAGAGGTGAG--GGAGATTGCTTGAACCCAGAGGTGGTGG 1700  
DB 16371 GTATCCCACTCTCAGAGGTTGAGGAGGAGATTGCTTGAACCCAGAGGTGGTGG 16430  
QY 1701 TTGACG----TGAGTTGAGCAATGCAATCCAGCTGGGCGAGAGGAGTGAATCTCT 1755  
DB 16431 TTGACGTTGAGCTGAGATCAATGCAATCCAGCTGGGCGAGAGGAGTGAATCTCTCT 16490  
QY 1756 CAAAAAAGAGATGCTCAACCTT 1788  
DB 16491 CTCTTAAAAAACAATCTTATAGCTATCTT 16523

## RESULT 5

US-09-402-532-38  
; Sequence 38, Application US/09402532  
; Patent No. 6498019  
; GENERAL INFORMATION:  
; APPLICANT: Taniyama, Yoshio  
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE  
; FILE REFERENCE: 2456USOP  
; CURRENT APPLICATION NUMBER: US/09/402,532  
; EARLIER APPLICATION NUMBER: PCT/JP98/01643  
; EARLIER FILING DATE: 1998-04-09  
; EARLIER APPLICATION NUMBER: JP 10-010289  
; EARLIER FILING DATE: 1998-01-22  
; EARLIER APPLICATION NUMBER: JP 9-184885  
; EARLIER FILING DATE: 1997-07-10  
; EARLIER APPLICATION NUMBER: JP 9-093355  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: P  
; SEQ ID NO 38  
; LENGTH: 2867  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Unsure (5' flanking promoter sequence of genomic DNA)  
US-09-402-532-38

Query Match  
Best Local Similarity 5.3%; Score 117.2; DB 4; Length 2867;  
Matches 195; Conservative 0; Mismatches 43; Indels 12; Gaps 4;

QY 1537 AGGCAAGTGAATCAAGAGAGTCAAGAGTTTGAACCACTGGCCCAATGAGAAAC 1596  
DB 104 AGGGGTGATGATCACTGAGAGTCAAGAGTTCAAGCAAGCCGTCACCAATGCGCAATC 163  
QY 1597 CCAATCTACCAAAAAATACAAATCAGCTGGCGGTGGT--GTGTGCTTGTAGTCCA 1654  
DB 164 CCGTCTCTCTAAAAAATACAAATCAGCTGGCGGTGGT--GTGTGCTTGTAGTCCA 1654  
QY 1655 ---ACGAGAGAGTTGAG--GGAATTTGCTTGAACCCAGAGAGTGGTGGCAATG 1709  
DB 224 GCAATTCAGAGAGCTGAGGAGAGAAATCGCTTGAACCCAGAGAGAGTGGCAATG 283  
QY 1710 ---GATTGACCAATCGAATCCAGCTGGGCGAGAGGAGTGAATCTCTCAAAAAA 1764

DB 284 GCCGGATACGCCCATCTGCTCCAGCTGGTGACAGCAAAACTGTGTCAAAAAA 343  
QY 1765 AAAAAAAGG 1774  
DB 344 AAAAAAAGG 353

## RESULT 6

US-09-810-671-3/C  
; Sequence 3, Application US/09810671  
; Patent No. 6455291  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000758  
; CURRENT APPLICATION NUMBER: US/09/810,671  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 21234  
; TYPE: DNA  
; ORGANISM: Human  
US-09-810-671-3

Query Match  
Best Local Similarity 5.3%; Score 117.2; DB 4; Length 21234;  
Matches 198; Conservative 0; Mismatches 48; Indels 12; Gaps 4;

QY 1529 GGGTGTGAGGCAAGTGAATCAAGAGGTCAGAGTTTGAACCACTGGCCCAATG 1588  
DB 11068 GGAGGCGAGGCAAGAGATCACTGAGGTGAGAGTTGAGACCAAGCTGGCTAATG 11009  
QY 1589 ATGAAACCCCATCTCTCAACAAAAATACAATAGCTGGCGGTGGTGGT--CCTG 1646  
DB 11008 GTGAACCCCATCTCTCAACAAAAATACAATAGCTGGCGGTGGTGGTGGTGGT 10949  
QY 1647 TAGTCCCA--ACGAGAGAGTTGAG--GGAGATTGCTTGAACCCAGAGGTGGTGT 1701  
DB 10948 TAATCCAGCTACTCAGAGGCTGAGTTTGAAGATCCGCGAATCTGGGAGTGGAGGT 10889  
QY 1702 TTGACG----GAGATTGAGCAATGCAATCCAGCTGGGCGAGAGTGAATCTCTC 1756  
DB 10888 TCGAGTGAAGCAGAGATGTACCACTGCACTCAGCTGGGTGAGAAAGTAACTCCATC 10829  
QY 1757 AAAAAAAGG 1774  
DB 10828 TCAAGAAAAAAGG 10811

## RESULT 7

US-10-109-854-3/C  
; Sequence 3, Application US/10109854  
; Patent No. 6650337  
; GENERAL INFORMATION:  
; APPLICANT: YAN, Chunhua et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; FILE REFERENCE: CL000758DIV  
; CURRENT APPLICATION NUMBER: US/10/109,854  
; PRIOR FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: 60/227,470  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: 09/810,671  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 21234  
; TYPE: DNA

ORGANISM: Homo sapiens  
US-10-109-854-3

Query Match 5.3%; Score 117.2; DB 4; Length 21234;  
Best Local Similarity 76.7%; Pred. No. 3.3e-26;  
Matches 198; Conservative 0; Mismatches 48; Indels 12; Gaps 4;

QY 1529 GGGGTGTGAGCAAGTATCATCAGAGTGTGAGAGTTCAGACCACTGGCCCAATG 1588  
DB 11068 GGAAGCCGAGGAGGAGAGATCCTGTGAGTTCAGAGTTCAGACCACTGGCCCAATG 11009  
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCCGTCTGTGTG--CCTG 1646  
DB 11008 GTGAACCCCATCTCTCTAAAACTATAAAAAATTACCCGGGTGTGTGTGGCCCACTG 10949  
QY 1647 TAGTCCCA---AGCAGAGAGTTGAG--GGGAGAAATGCTTGAACCCAGAGTGTGTG 1701  
DB 10948 TAATCCCAAGTACTCAGAGAGGCTGAGTTAGAGAAATGCGCGAAGCTG3AGGTGTGAGGT 10889  
QY 1702 TGCAGT-----GAGATTGAGCAACTGCATTCAGCTGGGCGAGCGAGTGAAGTGTCTC 1756  
DB 10888 TGCAGTGAAGCCGAGATGTCTCCACTGCACTCAGCTGGGTGACAAAGTGAATCCATC 10829  
QY 1757 AAAAAAAAAAAAAAAAAAG 1774  
DB 10828 TCAAGAGAAAAAAAAAAG 10811

RESULT 8  
US-09-435-739-42  
; Sequence 42; Application US/09435739  
; Patent No. 6664105  
; GENERAL INFORMATION:  
; APPLICANT: Beckman, Inc.  
; APPLICANT: Viodevsky, Israel  
; APPLICANT: Feinstein, Elena  
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY  
; TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS  
; FILE REFERENCE: 00/20454  
; CURRENT APPLICATION NUMBER: US/09/435,739  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: Patent version 3.0  
; SEQ ID NO 42  
; LENGTH: 44848  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-435-739-42

Query Match 5.3%; Score 117; DB 4; Length 44848;  
Best Local Similarity 73.1%; Pred. No. 6.7e-26;  
Matches 179; Conservative 0; Mismatches 60; Indels 6; Gaps 2;  
QY 1529 GGGGTGTGAGCAAGTATCATCAGAGTGTGAGAGTTCAGACCACTGGCCCAATG 1588  
DB 42887 GGAAGCCGAGGAGGAGATGATCCTGTGAGTTCAGAGAAATTGACACCACTGGCCCAATG 42946  
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCCGTCTGTGTGTGCTGTG 1648  
DB 42947 GTGAACCCCATCTCTCTAAAAATCAAAAAATTAGCTGGGTGTGTGTGTGTGTGTG 43006  
QY 1649 GTCCCAAGCAGAGAGTGTGAGGAGAAATGCTTGAACCCAGAGAGTGTGTGTGTGAGT- 1707  
DB 43007 GCTACTTGGAGGCTG--GGCAGAGAGAAATTAATTGAAACCCAGAGAGGAGTTGAGTGTG 43065  
QY 1708 ---GAGATTGAGCAACTGCATTCAGCTGGCCGAGCGAGTGAAGTGTCTCAAAAAAAA 1763  
DB 43066 AGCCGAGATGTGTGCACTGTCACTCCAGCTGGGTGTGAGAGAGAAATTCATCTCAAAAA 43125  
QY 1764 AAAAA 1768  
DB 43126 AACAA 43130

RESULT 9  
US-09-345-882-1/C  
; Sequence 1; Application US/09345882  
; Patent No. 639373  
; GENERAL INFORMATION:  
; APPLICANT: Bouguetel, Lydie  
; TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)  
; TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.  
; FILE REFERENCE: GENSER.031A  
; CURRENT APPLICATION NUMBER: US/09/345,882  
; PRIOR FILING DATE: 1999-06-30  
; PRIOR APPLICATION NUMBER: US 60/091,315  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 60/111,909  
; PRIOR FILING DATE: 1998-12-10  
; NUMBER OF SEQ ID NOS: 140  
; SOFTWARE: Patent.pm  
; SEQ ID NO 1  
; LENGTH: 162450  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 72794  
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 88073  
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 90842  
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 97122  
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 97152  
; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 99098  
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 99117  
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 103806  
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 106940  
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108106  
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108149  
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTT  
; FEATURE:  
; NAME/KEY: allele  
; LOCATION: 108308

```
OTHER INFORMATION: 5-135-357 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 108471
OTHER INFORMATION: 5-136-174 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134134
OTHER INFORMATION: 5-140-120 : polymorphic base C or T
FEATURE:
NAME/KEY: allele
LOCATION: 134362
OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 134374
OTHER INFORMATION: 5-140-361 : polymorphic base insertion of CA
FEATURE:
NAME/KEY: allele
LOCATION: 146328
OTHER INFORMATION: 5-143-84 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 146345
OTHER INFORMATION: 5-143-101 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 150329
OTHER INFORMATION: 5-145-24 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 160031
OTHER INFORMATION: 5-148-352 : polymorphic base G or T
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
FEATURE:
NAME/KEY: allele
LOCATION: 72771..72817
OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 8050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
FEATURE:
NAME/KEY: allele
LOCATION: 88050..88096
OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID52
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID49
FEATURE:
NAME/KEY: allele
LOCATION: 90819..90865
OTHER INFORMATION: complement polymorphic fragment 99-1437-325 SEQ ID70
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID32
FEATURE:
NAME/KEY: allele
LOCATION: 93690..93736
OTHER INFORMATION: polymorphic fragment 5-128-60 SEQ ID53
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID50
FEATURE:
NAME/KEY: allele
LOCATION: 97099..97145
OTHER INFORMATION: polymorphic fragment 99-1442-224 SEQ ID71
```

```
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
FEATURE:
NAME/KEY: allele
LOCATION: 97130..97177
OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
FEATURE:
NAME/KEY: allele
LOCATION: 99075..99121
OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
FEATURE:
NAME/KEY: allele
LOCATION: 99094..99140
OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
FEATURE:
NAME/KEY: allele
LOCATION: 103783..103828
OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
FEATURE:
NAME/KEY: allele
LOCATION: 106918..106966
OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
FEATURE:
NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:
```

```
Query Match 5.3%; Score 117; DB 4; Length 162450;
Best Local Similarity 76.4%; Pred. No. 1.7e-25;
Matches 198; Conservative 0; Mismatches 50; Indels 11; Gaps 4;

QY 1529 GGGTGTGAGGCAAGTATCAGAGAGTTCAGAGTTTGAGACCAACTGCGCAACATG 1588
Db 86346 GAGGCGAAGGAGGATGATCCTGAGGTGAGAGTTCAGAGCAGCTGGCCAAATG 86287
QY 1589 ATGAAACCCATCTTACCAAAAAAATACAAATCAGCTGCGC-CGTGCTGTGCTGT 1647
Db 86286 GTGAAACCTGCTCTACTTAAAAAATACAAATATGACGAGCATGTGGCGGCTGT 86227
QY 1648 AGTCCCA--ACGAGGAGGTTGAGG--GGGAAATGCTTGAACCCAGAGAGTGTGTT 1702
Db 86226 AATCCCACTCTCTGAGGCTGAGGAGGAGGAAATGCTTGAACCCGAGAGGCGAGGTT 86167
```



```

Oy      1703 GCAGTGAG-----ATTGAGCAACTGCAATCCAGCTGGGGGACGAGTGAAGTGTCTCA 1757
        |||
Db      86166 GCAGTGAGCCGATATTCGACCACTGCTCACTCAGCTTGGCGACAGAGTGAAGTCCGCTT 86107
Oy      1758 AAAAAAAAAAAAAAAAAAGAT 1776
        |||
Db      86106 CAAAAAAAAAAAAAAAAAGAT 86088

RESULT 10
US-09-759-359A-3/c
: Sequence 3, Application US/09759359A
: Patent No. 6492153
: GENERAL INFORMATION:
: APPLICANT: ABU-THREIDEN, Jane et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: FILE REFERENCE: CLO01043
: CURRENT APPLICATION NUMBER: US/09/759,359A
: CURRENT FILING DATE: 2001-01-16
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 90541
: TYPE: DNA
: ORGANISM: Human
US-09-759-359A-3

Query Match          5.3%; Score 116.8; DB 4; Length 90541;
Best Local Similarity 74.0%; Pred. No. 1,3e-25;
Matches 191; Conservative 0; Mismatches 57; Indels 10; Gaps 3;

Oy      1529 GGGTGTGAGGCAAGTAATCAAGAGGTCAGAGTTTGAAGACCACTGGCCAAACATG 1588
Db      5725 GGAGGCTGAGGCAAGAGAAATTACTTGTGAGTTCAGAGTTCAAGACCAAGCTGGCCAAACATG 5666
Oy      1589 ATGAAGCCCATCTCTCAACAAAAAATACAAATCAGCTGGC--CGTGTGTGTGCTGTG 1647
Db      5665 GTGAAGCCCATCTCTCTCTAATAAATACAAAGTTTAAACAGCCGTGTGTGATAGCTGTG 5606
Oy      1648 AGTCCCA---ACGAGAGAGGTTGAGGGAGAAATTGCTTGAACCCAGAGAGGTGTGTTC 1704
Db      5605 AGTTCAGCTACTCAGAGAGGTGCGACAGAGATTGCTTGAACCCAGAGAGGTGTGTTC 5546
Oy      1705 AGTGAGATTGAG-----CAATGCAATTCAGCTGGGCGACGAGTGAAGTCTCTCAA 1758
        |||
Db      5545 AGTGAACAGAGATCACACCACTGCACTAGACGCTGGGTGACAGAGCGAGATCTCATCTC 5486
Oy      1759 AAAAAAAAAAAAAAAAAAGAT 1776
        |||
Db      5485 AAAAAAAAAAAAAAAAAAT 5468

RESULT 11
US-09-918-686-1
: Sequence 1, Application US/09918686
: Patent No. 6475739
: GENERAL INFORMATION:
: APPLICANT: Brunkow, Mary
: APPLICANT: Prohl, Sean
: APPLICANT: Paepfer, Bryan
: APPLICANT: Staehling-Hampton, Karen
: TITLE OF INVENTION: METHODS FOR IDENTIFYING
: TITLE OF INVENTION: GENOMIC DELETIONS
: FILE REFERENCE: 240083.515
: CURRENT APPLICATION NUMBER: US/09/918,686
: CURRENT FILING DATE: 2001-07-30
: NUMBER OF SEQ ID NOS: 105
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 92139

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 7043, 8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

Query Match      5.3%; Score 116.8; DB 4; Length 92139;
Best Local Similarity 77.0%; Pred. No. 1.3e-25;
Matches 197; Conservative 0; Mismatches 47; Indels 12; Gaps 4;

QY      1529  GGGGTGTGGAGCAAGTAATACAGAGGTGAGAGTTTGAGCCAACTGGCCAACTG 1588
Db       85287  GGAAGCCGAGGCGAGGTGATCATCTGAGGCGAGAGTTGAGACCGACTGGCCAACTG 85346

QY      1589  ATGAAACCCCATCTCTTCCAAAAAAATTAACAATCAAGCTGGCCGTGTGG--TGTCCCTG 1646
Db       85347  GTGAACCTGCTCTCTTCACTACAAATACAAAAAATTGCTGGGTGTGTGGCAATACCTG 85406

QY      1647  TAGTCCCA---ACGCAAGAGTTGAGG--GGAAGATTGCTTGAACCCAGAGAGTGTGCT 1701
Db       85407  TAGTTCACGTATCTCAGGAAGCTGAGAGCGAGAAATTGCTTGAATCTCAGAGAGCAAGAT 85466

QY      1702  TGCAGTAGAG----ATTAGCAACTGCATTCACGCTGGGCGACGAGGTGAAGCTGTCTC 1756
Db       85467  TGCAGTAGGCAACATTTGGCCCATTCCTCCAGCCTGGGTGACAGAGCGAAGCTCTGTTC 85526

QY      1757  AAAAAAAAAAAAAAAAAA 1772
Db       85527  TCAAAAAAAAAAAAAA 85542

```

```

RESULT 12
US-09-800-960-3/c
; Sequence 3, Application US/09800960
; Patent No. 6387677
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1001158
; CURRENT APPLICATION NUMBER: US/09/800,960
; CURRENT FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 62804
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(62804)
; OTHER INFORMATION: n = A,T,C or G
US-09-800-960-3

```

	Query March Best Local Matches 179;	Similarity 78.2%; Conservative 0;	Score 116.2; Pred. No. 1.6e-25; Mismatch 43;	DB 4; Indels 7; Gaps 3;	Length 62804;
Oy	1551	CAGAGGTCGAGATTGAGACCAACCTGAGCAACATGATGAAACCCCATCTCTACCAA			1610
Db	29739	CATGAGATCAGGAGATCGAAACCATCTTAATCTAACGAGTGAAACCCCATCTCTACTAA			29680
Oy	1611	AAAAATCAAAATCAGCTGGCCGTCGTGTGTG--CCTTATGTCCTCA---ACGACAGAGGT			1665
Db	29679	AAATCAAAAATATTAGCCCGGCGGTGTGTGGCGGCACCTGTATGTCCTCCAGCTCTCGGAGGC			29620
Oy	1666	TGAGG--GGAGAAATTCCTTGAAACCCAGAGAGTGTTGACATGAGATGAGCAATGTC			1723
Db	29619	TGAGGCGAGGAATGCGCATGAAACCCGGAGGCGAGAGCTTGACATGAGACCGGCGCACTGC			29560

Qy 1724 AATCCAGCTGGGCGACGAGTGAAGTCTCTCAAAAAAAAAAAAAAAAA 1772  
Db 29559 ACTCCAGCTGGGCGACGAGTGAAGTCTCTCAAAAAAAAAAAAAAAAA 29511

## RESULT 13

US-10-096-960-3/c  
Sequence 3, Application US/10096960  
Patent No. 6664085  
GENERAL INFORMATION:  
APPLICANT: YE, Jane et al.  
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
FILE REFERENCE: CL001158DIIV  
CURRENT APPLICATION NUMBER: US/10/096,960  
PRIOR FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 09/800,960  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 62804  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)-(62804)  
OTHER INFORMATION: n = A,T,C or G  
US-10-096-960-3

Query Match 5.2%; Score 116.2; DB 4; Length 62804;  
Best Local Similarity 78.2%; Pred. No. 1.6e-25;  
Matches 179; Conservative 0; Mismatches 43; Indels 7; Gaps 3;

Qy 1551 CAGGAGTCGAGGAGTTGAGACCACTTGGCCCAAGATGATGAACCCCATCTACCAAA 1610  
Db 29739 CAGGAGTCGAGGAGTTGAGACCACTTGGCCCAAGATGATGAACCCCATCTACCAAA 29680  
Qy 1611 AAAAAATCAATCAGCTGCGCGCTGTGTG--CCTGTAGTCCCA--ACGACGAGGT 1665  
Db 29679 AATACAAAAAATTAAGCCGCGCTGTGTG--CCTGTAGTCCCA--ACGACGAGGT 29620  
Qy 1666 TGAGG--GGAGAAATTCCTTGAACCCAGAGGTGTGTGTTGAGAGATGAGCACTGC 1723  
Db 29619 TGAGGAGAGAGATGAGCAATGAACCCGAGGAGCAAGAGTTGAGAGAGACCGCGCACTGC 29560  
Qy 1724 AATCCAGCTGGGCGACGAGTGAAGTCTCTCAAAAAAAAAAAAAAAAA 1772  
Db 29559 ACTCCAGCTGGGCGACGAGTGAAGTCTCTCAAAAAAAAAAAAAAAAA 29511

## RESULT 14

US-09-820-005-3/c  
Sequence 3, Application US/09820005  
Patent No. 6489149  
GENERAL INFORMATION:  
APPLICANT: SHAO, Wei et al  
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
FILE REFERENCE: CL001198  
CURRENT APPLICATION NUMBER: US/09/820,005  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 18853  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)....(18853)

OTHER INFORMATION: n = A,T,C or G  
US-09-820-005-3

Query Match 5.2%; Score 115.8; DB 4; Length 18853;  
Best Local Similarity 77.9%; Pred. No. 8.6e-26;  
Matches 194; Conservative 0; Mismatches 42; Indels 13; Gaps 4;

Qy 1537 AGCGAAGTGAATCAAGAGGTGAGAGTTGAGACCACTTGGCCCAAGATGATGAAC 1596  
Db 12720 AGCGGAGGAGTCACTGAGGTGAGAGTTGAGACCACTTGGCCCAAGATGATGAAC 12661  
Qy 1597 CCATCTCAACCAAAAAAATCAATCAGTGTGCTGTGTG--TGCTGTAGTCCCA 1654  
Db 12660 CCATCTCACTAAATAATCAAAAAAATTAAGCTGTGCTGTGTGCTGTGTGCTGTG 12601  
Qy 1655 ---ACGACGAGGTGAGG--GGAGAAATTCCTTGAACCCAGAGGTGTGTGCTGAC-- 1706  
Db 12600 GGTACTCAGGTGTGAGGAGAGAAATTCCTTGAACCCAGAGGTGTGTGCTGAC-- 12541  
Qy 1707 ---TGAGATTGAGCACTGCAATCCAGCTGGGCGACGAGTGAAGTGTCTCAAAAAA 1763  
Db 12540 AGCTGAGATTGCAATTACTGCACTCCAGCTGGGTGACAGACCAAGATTGTCTCAAAA 12481  
Qy 1764 AAAAAAAAA 1772  
Db 12480 AAAAAAAAA 12472

## RESULT 15

US-09-851-896-11  
Sequence 11, Application US/09851896  
Patent No. 6410325  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Freier  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2--INDEPENDENT  
FILE REFERENCE: R15-0220  
CURRENT APPLICATION NUMBER: US/09/851,896  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 11  
LENGTH: 2296  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-851-896-11

Query Match 5.2%; Score 115.4; DB 4; Length 2296;  
Best Local Similarity 73.9%; Pred. No. 2.4e-26;  
Matches 190; Conservative 0; Mismatches 56; Indels 11; Gaps 3;

Qy 1527 CTGGGCTGTGAGGCAAGTGAATCAAGAGGTGAGAGTTGAGACCACTTGGCCCAACA 1586  
Db 246 CGGAGGCTGAGGCGGAGATCACTTGAAGTCAAGAGTTGAGACCACTTGGCCCAACA 305  
Qy 1587 TGATGAACCCCATCTTACCAAAAAAATCAATCAGTGTGCTGTGTGCTGTGCTGCTG 1643  
Db 306 TGATGAACCCCATCTTACCAAAAAAATCAATCAGTGTGCTGTGTGCTGTGCTGCTG 365  
Qy 1644 --CTGTAGTCCCAACGAGAGGTTGAG--GGAGAAATTCCTTGAACCCAGAGAGGTGTG 1699  
Db 366 TGCAATCTCACTGAGGAGGAGGCTGACGAGGAAATCTGTGAATCCGAGAGGCGGAG 425  
Qy 1700 GTTGCACT---GAGATTGAGCACTGCAATCCAGCTGGGCGACGAGTGAAGTGTCT 1755  
Db 426 GTTGCACTGAGGAGAGTGTGCTGCACTGCACTGCACTGCACTGCACTGCACTGCACTG 485  
Qy 1756 CAAAAAAAAAAAAAAAA 1772  
Db 486 TCAAAAAAAAAAAAAAAAA 502

Mon Mar 29 09:56:41 2004

us-09-622-964-1\_copy\_13900\_16115.rn1

Page 9

Search completed: March 25, 2004, 07:24:26  
Job time : 170 secs

---

*This Page Blank (uspto)*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 25, 2004, 05:47:23 ; Search time 1271 Seconds  
(without alignments)  
6491.101 Million cell updates/sec

Title: US-09-622-964-1\_COPY\_13900\_16115

Perfect score: 2216  
Sequence: 1 aactctgtcttcttcacgc.....cctatgagtttaataac 2216

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing First 45 summaries

Database :

Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US05B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1631.6	73.6	1717	15	US-10-264-237-259 Sequence 259, App
2	643.8	29.1	1263	10	US-09-746-783-3 Sequence 3, Appl1
3	404.2	18.2	459	10	US-09-918-995-26929 Sequence 26929, A
4	339.4	15.3	1326	13	US-10-071-766-10 Sequence 10, Appl
5	164	7.4	1198	9	US-09-880-107-2174 Sequence 2174, Ap
6	133.2	6.0	747	15	US-10-027-632-173516 Sequence 173516,
7	133.2	6.0	65042	14	US-10-229-124-3 Sequence 3, Appl1
8	133.2	6.0	149480	10	US-09-873-367C-284 Sequence 284, App
9	133.2	6.0	149480	10	US-09-873-367C-285 Sequence 285, Appl
10	127.6	5.8	322101	14	US-10-060-902-1 Sequence 1, Appl1
11	126.2	5.7	22645	10	US-09-764-891-7673 Sequence 7673, Ap
12	126.2	5.7	22645	10	US-09-764-891-8183 Sequence 8183, Ap
13	124.2	5.6	39000	10	US-09-957-956-5 Sequence 5, Appl1
14	124.2	5.6	130320	15	US-10-408-168-1 Sequence 1, Appl1
15	123.8	5.6	14874	10	US-09-764-891-7672 Sequence 7672, Ap

16	123.4	5.6	753	15	US-10-027-632-21017 Sequence 21017, A
17	123.4	5.6	10500	14	US-10-312-495-13 Sequence 13, Appl
18	123.4	5.6	358246	15	US-10-292-798-1095 Sequence 1095, Ap
19	123.2	5.6	52216	9	US-09-747-810-1 Sequence 1, Appl1
20	123	5.6	7720	9	US-09-954-456-946 Sequence 946, App
21	123	5.6	7720	9	US-09-954-456-1589 Sequence 1589, Ap
22	122.8	5.5	433	15	US-10-027-632-256214 Sequence 256214,
23	122.4	5.5	1094	15	US-10-027-632-265287 Sequence 265287,
24	122.2	5.5	31883	15	US-10-027-632-256666 Sequence 256666,
25	121.8	5.5	23748	10	US-09-764-891-7917 Sequence 7917, Ap
26	121.8	5.5	39703	14	US-10-017-161-1041 Sequence 1041, Ap
27	121.8	5.5	39703	15	US-10-292-798-883 Sequence 883, App
28	121.8	5.5	134292	12	US-10-240-425-1102 Sequence 1102, Ap
29	121.6	5.5	36159	13	US-10-135-687-3 Sequence 3, Appl1
30	121.4	5.5	723	15	US-10-027-632-14278 Sequence 14278, A
31	121.4	5.5	733	15	US-10-027-632-14279 Sequence 14279, A
32	121.2	5.5	433	15	US-10-027-632-256213 Sequence 256213,
33	121.2	5.5	11282	9	US-09-764-847-1747 Sequence 1747, Ap
34	121.2	5.5	11282	14	US-10-092-154-1747 Sequence 1747, Ap
35	121.2	5.5	13630	9	US-09-764-868-1369 Sequence 1369, Ap
36	121.2	5.5	13630	11	US-09-764-875-1320 Sequence 1320, Ap
37	121.2	5.5	21221	14	US-10-017-161-2067 Sequence 2067, Ap
38	121	5.5	761	15	US-10-027-632-25603 Sequence 25603, A
39	121	5.5	21721	9	US-09-764-853-861 Sequence 861, App
40	121	5.5	162025	14	US-10-272-665-35 Sequence 35, Appl
41	121	5.5	162025	14	US-10-272-665-36 Sequence 36, Appl
42	121	5.5	162025	14	US-10-273-321-35 Sequence 35, Appl
43	121	5.5	162025	14	US-10-273-321-36 Sequence 36, Appl
44	121	5.5	162025	14	US-10-272-756-35 Sequence 35, Appl
45	121	5.5	162025	14	US-10-272-756-36 Sequence 36, Appl

## ALIGNMENTS

RESULT 1  
US-10-264-237-259  
; Sequence 259, Application US/10264237  
; Publication No. US20040009491A1  
; GENERAL INFORMATION:  
; APPLICANT: Btise et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P4131P1  
; CURRENT APPLICATION NUMBER: US/10/264,237  
; CURRENT FILING DATE: 2002-10-04  
; PRIOR APPLICATION NUMBER: PCT/US01/16450  
; PRIOR FILING DATE: 2001-05-18  
; PRIOR APPLICATION NUMBER: US 60/205,515  
; PRIOR FILING DATE: 2000-05-19  
; NUMBER OF SEQ ID NOS: 2876  
; SOFTWARE: PatentIn Ver. 3.1  
; SEQ ID NO 259  
; LENGTH: 1717  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1072)..(1072)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1596)..(1596)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1634)..(1634)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1657)..(1657)  
; OTHER INFORMATION: n equals a,t,g, or c  
; FEATURE:  
; NAME/KEY: misc\_feature

	Pred.	No. 0;	Mismatches	Indels	Gaps
Matches 1692; Conservatively	98.5%;	3	18;	5;	5;

Db	900	TGACAGCCAGGCGCTTGAAGCTGAGACAGATGTTATCACTGACCCCAACTTACTTTAGCAAGG	959
Qy	1017	GTGGCTGACCCCAAMACATGAGTGGCAGTCAAGCTGATGACATGTAAACAATCTTCCCCCA	107
Db	960	GTGGCTGACCCCAAMACATGAGTGGCAGTCAAGCTGATGACATGTAAACAATCTTCCCCCA	101
Qy	1077	TAACTATTAGGGTGTGTGCCAAGACATACAGAAAGGTGGAGAACTGCTCTCACTCC	1133
Db	1020	TAACTATTAGGGTGTGTGCCAAGACATACAGAAAGGTGGAGAACTGCTCTCACTCC	1076
Qy	1137	TAGGAACCTGTATGATGTGAGGTTGAAGGGTGTCAAGGCCCTTAAGTCAATTTCTCACTG	1196
Db	1079	TAGGAACCTGTATGATGTGAGGTTGAAGGGTGTCAAGGCCCTTAAGTCAATTTCTCACTG	1138
Qy	1197	CCTGGGAACCTACCAAAATACCTCTTGCTCTCTTGGGGTCAAGCCCAAGCTGTACAAA	1256
Db	1139	CCTGGGAACCTACCAAAATACCTCTTGCTCTCTTGGGGTCAAGCCCAAGCTGTACAAA	1198
Qy	1257	ATCAGATATTTCCCTTTATTTCCAGATTTCTGAGACCTGTCAACCAATATTAACACCCC	1316
Db	1199	ATCAGATATTTCCCTTTATTTCCAGATTTCTGAGACCTGTCAACCAATATTAACACCCC	1258
Qy	1317	ACTTCAGCCCCCAATCAAGCTGGAGAGGTAACTTCCCTTTCTGGAATTTCTCAAGAGT	1376
Db	1259	ACTTCAGCCCCCAATCAAGCTGGAGAGGTAACTTCCCTTTCTGGAATTTCTCAAGAGT	1318
Qy	1377	TACTTCAAGGGTCAAGAACAGCACTATATGATTAAGAACCTTAAAGGCAACAATTT	1436
Db	1319	TACTTCAAGGGTCAAGAACAGCACTATATGATTAAGAACCTTAAAGGCAACAATTT	1377
Qy	1437	CANCTTGTCTTCAAGCTTAAGACAGAACTTGGCAAAATCTGTGGCCTGTTCAGAAAG	1496
Db	1378	CANCTTGTCTTCAAGCTTAAAGCAGGACCTTGGCAAAATCTGTGGCCTGTTCAGAAAG	1437
Qy	1497	GATGTCATATTTAAAGATCTTGTCTTGGGCTGGGTGTGAAGCAAGTGAATCAAGAG	1556
Db	1438	GATGTCATATTTAAAGATCTTGTCTTGGGCTGGGTGTGAAGCAAGTGAATCAAGAG	1497
Qy	1557	GTCAAGAGATTGAAGCAACCTGGCCAAATGATTAAGAACCCCATCTTCAACAAAAAAT	1616
Db	1498	GTCAAGAGATTGAAGCAACCTGGCCAAATGATTAAGAACCCCATCTTCAACAAAAAAT	1557
Qy	1617	AATAATCAGCTGCGCGTGTGTGTGTGCTGTATGCCCAAGCAGAGGTTAGGGGAGAA	1676
Db	1558	AATAATCAGCTGCGCGTGTGTGTGTGCTGTATGCCCAAGCAGAGGTTAGGGGAGAA	1617
Qy	1677	TTGCTTGAACCAAGAGTGTGTGTGCAAGTGAATGA-GCAATGCAATCAAGCCTGG	1735
Db	1618	TTGCTTGAACCAAGAGTGTGTGTGCAAGTGAATGAAGTCAATCAAGCCTGG	1677
Qy	1736	GCGAC-GGAGTGAAGCTGTCTCAAAAAA	1772
Db	1678	GCGACGGAGGTGAAGCTGTCTCAAAAAA	1715

CORRESPONDENCE ADDRESS:  
ADDRESS: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: MA  
COUNTRY: U.S.A.  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/746,783  
FILING DATE: 21-Dec-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Milasincic, Debra J.  
REGISTRATION NUMBER: 46,931  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1263 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-746-783-3

Query Match 29.1%; Score 643.8; DB 10; Length 1263;  
Best Local Similarity 99.7%; Pred. No. 1.7e-196;  
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

17 CAGCCGGAACAAAGAGAGATGAGTTCAGCCCAATCAGGAGAGAGAGATGCTCA 76  
225 CAGCCTGAACAAAGAGAGATGAGTTCAGCCCAATCAGGAGAGAGAGATGCTCA 284  
77 CGCTGACATCATTTGGCGCTTCTTCAAGGCTGAGTCCATGATCACCATTCTCCAGG 136  
285 CGCTGACATCATTTGGCGCTTCTTCAAGGCTGAGTCCATGATCACCATTCTCCAGG 344  
137 AAACCTCAAGAGCAAACTAAGTGGCCCAAGAGAGAAATCCCTTCTCAAGAGGCTG 196  
345 AAACCTCAAGAGCAAACTAAGTGGCCCAAGAGAGAAATCCCTTCTCAAGAGGCTG 404  
197 CAAAAACCAAG 256  
405 CAAAAACCAAG 464  
257 GCTTAAAGGCTGAGAGGCTTCAAGTCTGCCCCCACTGATCAGAGGCCAGGCTACTA 316  
465 GCTTAAAGGCTGAGAGGCTTCAAGTCTGCCCCCACTGATCAGAGGCCAGGCTACTA 524  
317 TGGCCCAAGAGAGGCTTCAAGTCTGCCCCCACTGATCAGAGGCCAGGCTACTA 376  
525 TGGCCCAAGAGAGGCTTCAAGTCTGCCCCCACTGATCAGAGGCCAGGCTACTA 584  
377 GTCAAGAGCTTCAAGTGTCAAGGAGATAGACCAAGAGAGAGAGAGAGAGAGAG 436  
585 GTCAAGAGCTTCAAGTGTCAAGGAGATAGACCAAGAGAGAGAGAGAGAGAGAG 644  
437 TTCTGGGGGCAAG 496  
645 TTCTGGGGGCAAG 704  
497 CCCAGAGATATCTCAAGTGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556  
705 CCCAGAGATATCTCAAGTGTCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 764  
557 GATCCCGGAG 616

Db 765 GATCCCGGAG 824  
Qy 617 ACTCAAGATCATATGATCTTATTTGGGCTTTGAGAGAGAGAGAGAGAGAG 663  
Db 825 ACTCAAGATCATATGATCTTATTTGGGCTTTGAGAGAGAGAGAGAGAGAG 871

RESULT 3  
US-09-918-995-26929  
Sequence 26929, Application US/09918995  
Publication No. US20030073623A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
FILE REFERENCE: 20411-756  
CURRENT APPLICATION NUMBER: US/09/918,995  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: US/09/235,076  
PRIOR FILING DATE: 1999-01-20  
NUMBER OF SEQ ID NOS: 38054  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 26929  
LENGTH: 459  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(459)  
OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-26929

Query Match 18.2%; Score 404.2; DB 10; Length 459;  
Best Local Similarity 96.4%; Pred. No. 2.3e-119;  
Matches 423; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

Qy 1333 CGTGGAGAGAGATGATTAATCTCCCTTTCTGAGATTCAGAGCACTTTCAGGGTCAG 1392  
Db 22 CGAGGAGATTCGTAATCTCCCTTTCTGAGATTCAGAGCACTTTCAGAGCACTTTCAG 80  
Qy 1393 AACAGCAGCTATTAATGATTAAGAACTTAAAGGAGAGAGAGAGAGAGAGAGAG 1452  
Db 81 AACAGCAGCTATTAATGATTAAGAACTTAAAGGAGAGAGAGAGAGAGAGAGAG 140  
Qy 1453 CTAAAGCAGAGAACTTGGCAAACTGTCGCTTTGAGCAAGAGAGAGAGAGAG 1512  
Db 141 GCTAAGCAGAGAACTTGGCAAACTGTCGCTTTGAGCAAGAGAGAGAGAGAGAG 200  
Qy 1513 AATCTTGTCTTGGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1572  
Db 201 AATCTTGTCTTGGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 260  
Qy 1573 CAACCTGGCAGACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1632  
Db 261 CAACCTGGCAGACATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 320  
Qy 1633 TGTGTGTGTCTGTAGTCCCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1692  
Db 321 TGTGTGTGTCTGTAGTCCCAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380  
Qy 1693 GGTGTGTGTGAG 1752  
Db 381 GGTGTGTGTGAG 440  
Qy 1753 TCTCAAAAAAAAAAAAAA 1771  
Db 441 TCTCAAAAAAAAAAAAAA 459

RESULT 4  
US-10-071-766-10/c  
Sequence 10, Application US/10071766

Publication No. US20020192678A1  
GENERAL INFORMATION:  
APPLICANT: Huel-Mei Chen  
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE  
FILE REFERENCE: PA-0043 US  
CURRENT APPLICATION NUMBER: US/10/071,766  
CURRENT FILING DATE: 2002-02-07  
NUMBER OF SEQ ID NOS: 144  
SOFTWARE: PERL Program  
SEQ ID NO 10  
LENGTH: 1326  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2  
US-10-071-766-10

Query Match 15.3%; Score 339.4; DB 13; Length 1326;  
Best Local Similarity 99.7%; Pred. No. 4.2e-98;  
Matches 340; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1876 AGCAGATTCCTTAACCTGCTTCTTAATGGGATGCTTGCCAGCAGGCTCTCACTGTGT 1935  
DB 1332 AGCAGATTCCTTAACCTGCTTCTTAATGGGATGCTTGCCAGCAGGCTCTCACTGTGT 1267  
QY 1936 GTACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1995  
DB 1266 GTACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1207  
QY 1996 TCCACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2055  
DB 1206 TCCACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1147  
QY 2056 CTTTAAATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2115  
DB 1146 CTTTAAATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1087  
QY 2116 AACTCAGACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 2175  
DB 1086 AACTCAGACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 1027  
QY 2176 CACTTAAATGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 2216  
DB 1026 CACTTAAATGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 986

## RESULT 5

US-09-880-107-2174/c  
Sequence 2174; Application US/09880107  
Patent No. US20020142981A1  
GENERAL INFORMATION:  
APPLICANT: Horne, Darci T.  
APPLICANT: Vockley, Joseph G.  
APPLICANT: Scherf, Iwe  
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer  
FILE REFERENCE: 44921-5028-WO  
CURRENT APPLICATION NUMBER: US/09/880,107  
PRIOR FILING DATE: 2001-06-14  
PRIOR APPLICATION NUMBER: US 60/211,379  
PRIOR FILING DATE: 2000-06-14  
PRIOR APPLICATION NUMBER: US 60/237,054  
NUMBER OF SEQ ID NOS: 3950  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2174  
LENGTH: 1198  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941  
US-09-880-107-2174

Query Match 7.4%; Score 164; DB 9; Length 1198;  
Best Local Similarity 100.0%; Pred. No. 1.9e-41;  
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2053 CAGCCTTAAATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2112  
DB 1198 CAGCCTTAAATGCTTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1139  
QY 2113 TTTAACTCAGACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 2172  
DB 1138 TTTAACTCAGACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 1079  
QY 2173 CCACACTTAAATGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 2216  
DB 1078 CCACACTTAAATGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAG 1035

## RESULT 6

US-10-027-632-173516/c  
Sequence 173516; Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 173516  
LENGTH: 747  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-173516

Query Match 6.0%; Score 133.2; DB 15; Length 747;  
Best Local Similarity 82.8%; Pred. No. 1.2e-31;  
Matches 202; Conservative 0; Mismatches 33; Indels 9; Gaps 4;

QY 1537 AGCAGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAA 1596  
DB 454 AGCAGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAA 395  
QY 1597 CCATCTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 1654  
DB 394 CCATCTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 336  
QY 1655 ---AGCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 1708  
DB 335 GCTACTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 276  
QY 1709 AGATTGAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 1768  
DB 275 AGATTGAGAGTGGGAACTCTGATTCAGAGTGGGAACTCTGATTCAGAGTGGGAACTCT 216  
QY 1769 AAAA 1772



Db 215 AAAA 212

RESULT 7  
US-10-229-124-3/C  
Sequence 3, Application US/10229124  
Publication No. US2003003168A1  
GENERAL INFORMATION:  
APPLICANT: WEI, Ming-Hui et al.  
TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
FILE REFERENCE: CL001139DIV  
CURRENT APPLICATION NUMBER: US/10/229,124  
CURRENT FILING DATE: 2002-08-28  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 65042  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(65042)  
OTHER INFORMATION: n = A,T,C or G  
US-10-229-124-3

Query Match 6.0%; Score 133.2; DB 14; Length 65042;  
Best Local Similarity 70.6%; Pred. No. 2,6e-30;  
Matches 209; Conservative 0; Mismatches 78; Indels 9; Gaps 2;

QY 1527 CTGGGTGTGGAGCGAAGTGAATCAAGAGGTGAGAGTTGAGACCACTGGCCACACA 1586  
Db 35972 CGGAGAGCCGAGGCGGCGGATCACTGAGGTCCAGACCAAGCTGGCCACACA 35913

QY 1587 TGATGAAACCCCATCTCTACCAAAAAAATACAAATCAGTCGCGCTGTGCTGCTG 1646  
Db 35912 TGGTGAACCCCATCTCTACCAAAAAAATACAAATCAGTCGCGCTGTGCTGCTG 35853

QY 1647 TAGTCCCAAC---GCAGAGGTTGAGGGAGAAATTCCTTGAACCCAGAGGTGTGTT 1702  
Db 35852 TAATCCCAAGTACTCAAGAGGCTGGGCGAGCAATCACTTGAACCCAGAGGTGTGAGTT 35793

QY 1703 GCAG-----TGAGATTGAGCACTGCAATCCAGCTGGGCGAGAGTGAAGTGTCTCA 1757  
Db 35792 GCGGTGAGCTGAGATTGTGCACCTGCACTGAGCTGGGCGAAGAGAGGAGCTGTCT 35733

QY 1758 AAAAAAAAAAAGGATCGTCTCAACTTTGCCCTCTACTGCAACATTTTGT 1813  
Db 35732 CAAAAAAAAAAAAAAAAAAAAAAAAAAGTTATGCTGTGAGCTAAACATTCAGT 35677

RESULT 8  
US-09-873-367C-284  
Sequence 284, Application US/09873367C  
Publication No. US20030165839A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
APPLICANT: Soppet, Daniel  
APPLICANT: Andrews, Gregory  
APPLICANT: Augustus, Meena  
APPLICANT: Ebner, Reinhard  
APPLICANT: Carter, Kenneth  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
Signature Gene Sets  
FILE REFERENCE: 689290-64  
CURRENT APPLICATION NUMBER: US/09/873,367C  
CURRENT FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: U.S. 60/236,891  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: U.S. 60/236,842  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: U.S. 60/244,867

QY 1529 GGGTGTGAGCGCAAGTATCAAGAGGTGAGAGTTTGAAGCAACTGCGCAACATG 1588  
Db 8803 GGAGGCTGAGGCGAGTGATCACTTGAGGTCAAGAGTTTAAAGACAGCTGGCCACATG 8862

QY 1589 ATGAAACCCCATCTACCAAAAAAATCAATCAGCTGGCGGTGTGTTG---TGCGTG 1646  
Db 8863 GTAAACCCCATCTCTACCAAAAAAATTTAGCTGGGTGTGTGTGATGCTGCTG 8922

QY 1647 TAGTCCCAACGACGAGG---TTGAGGGAGAAATGCTTGAACCCAGAGGTGTGTG 1701  
Db 8923 TAGTCCCAAGTACTGGGAGGCTGAGACAGAGAAATTTGCTTGAATCAGAGGTGTGAGT 8982

QY 1702 TGCAGT---GAGATTGAGCAATCGAATCCAGCTGGGCGACGAGTGAAGTGTCTC 1756  
Db 8983 TGCAGTAAGCGGAGTCTGCTCCATTCACCTCGGCTGTGACAGCGAGACTGTCTC 9042

QY 1757 AAAAAAAAAAAAAA 1772  
Db 9043 TCAAAAAAAAAAAAAA 9058

RESULT 9  
US-09-873-367C-285  
Sequence 285, Application US/09873367C  
Publication No. US20030165839A1  
GENERAL INFORMATION:  
APPLICANT: Young, Paul  
APPLICANT: Soppet, Daniel  
APPLICANT: Andrews, Gregory  
APPLICANT: Augustus, Meena  
APPLICANT: Ebner, Reinhard  
APPLICANT: Carter, Kenneth  
TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using  
Signature Gene Sets  
FILE REFERENCE: 689290-64  
CURRENT APPLICATION NUMBER: US/09/873,367C  
CURRENT FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: U.S. 60/236,891  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: U.S. 60/236,842  
PRIOR FILING DATE: 2000-09-29  
PRIOR APPLICATION NUMBER: U.S. 60/244,867  
PRIOR FILING DATE: 2000-11-01  
PRIOR APPLICATION NUMBER: U.S. 60/245,084  
PRIOR FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 1067  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 285  
LENGTH: 149480  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-873-367C-285

Query Match 6.0%; Score 133.2; DB 10; Length 149480;  
Best Local Similarity 78.5%; Pred. No. 4,7e-30;  
Matches 201; Conservative 0; Mismatches 43; Indels 12; Gaps 3;

QY 1529 GGGTGTGAGCGCAAGTATCAAGAGGTGAGAGTTTGAAGCAACTGCGCAACATG 1588  
Matches 201; Conservative 0; Mismatches 43; Indels 12; Gaps 3;

RESULT 11  
 US-09-764-891-7673/c  
 Sequence 7673, Application US/09764891  
 Publication No. US20030077808A1  
 GENERAL INFORMATION  
 APPLICANT: Rosen et al.  
 TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
 FILE REFERENCE: PC06  
 CURRENT APPLICATION NUMBER: US/09/764, 891

Query Match	5.7*	Score 126.2	DB 10	Length 22645
Best Local Similarity	78.8*	Pred. No. 2.3e-26		
Matches	204	Conservative 0	Mismatches 43	Indels 12
			Gaps 4	
1529	GGGTGTGAGCAGAGTAATCACAGAGGTTCAGAGATTGAGCAACCTGGCCACATG			1588
11375	GGAGGCTGAGGCGAGCGGATCACTTGGAGTTCAGAGATTTCAGAGATCACTGGCCACATG			11316
1589	ATGAACCCCATCTCTTCACAAAAAATATCAATCACTGGCGCTGCTGTG--TGCTTG			1646
11315	GTGAACCCCTGTCTCTACAAAAATACAAAAATATATTAAGTGTGTGGTCATGCTTG			11256
1647	TAGTCCCA---ACGACAGAGTTGAG--GGAGATATGCTTGAACCCAGAGGTGATGAT			1701
11255	TAAATCCAGACTTGGAGGCTGAGGTGGGAGATCATTTGAACCCAGAGGTGAGAT			11196
1702	TGACAGT----AGATTGAGCACTGCAATCCAGCGCTGGCGGACGAGATGAGACTGTCTC			1756
11195	TGCATATGCTAAAGATTGTGCTCATTTGACATTCACGCTGGGGCAACAGAGGGAGACTGTTTC			11136

QY	1757	AAAAAAAAAAAAAAAAAGGA	1775
Db	11135	AAAAAAAAAAAAAAAAAGAA	11117

## RESULT 13

```

US-09-957-956-5
/ Sequence 5, Application US/09957956
/ Publication No. US20030130215A1
/ GENERAL INFORMATION:
/ Applicant: Ryan, James W.
/ TITLE OF INVENTION: Genomic Polynucleotide Fragments From Chromosome 7
/ FILE REFERENCE: JR-14,000-US
/ CURRENT APPLICATION NUMBER: US/09/957,956
/ CURRENT FILING DATE: 2001-09-21
/ PRIOR APPLICATION NUMBER: US 60/234,422
/ PRIOR FILING DATE: 2001-09-21
/ NUMBER OF SEQ ID NOS: 8
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 5
/ LENGTH: 39000
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-957-956-5

```

Query Match	5.6%	Score 124.2	DB 10	Length 39000
Best Local Similarity	76.3%	Pred. No. 1.5e-27		
Matches 196; Conservative	0	Mismatches 48	Indels 13	Gaps 3

QY	1529	GGGGTGTGAGCAAGTCAATTCACAGGAGGTCAAGAGTTTGAAGCAAACTGGGCAACATG	1588
Db	36491	GGAGGCTGAGGACAGTATGATCATCCTGAGTCTCAGAGTTGAAGACCAAGCTGGCCACACAG	36550
QY	1589	ATGAACCCCACTCTTACCAAAAAAATACAAATCAGCTGGCCGTGCTGT- GTGCTTG	1646
Db	36551	GTGAACCCCACTCTCTCTAAAAAATCAAAAAATAGCTGGCGGTGTGTGTGGTCTCTG	36610
QY	1647	TAGTCCCAAC-----GCAGAGGTTGAGGGGGAATGTGCTTAACCCAGAGGTGGTGG	1700
Db	36611	TNATCCAGCTACTCGGAGGCTGAGGCAAGGAAATGTGCTTAACCCAGAGGCCGAGG	36670
QY	1701	TTGCAGTGAG----ATTGAGCAATCGCAATCCAGCTGGGCGACGAGATGAGACTGTCT	1755
Db	36671	TTGCAGTGAGCAAAAATCACAACCACTGCACTCCAGCTGGGTGACAGATGAGATTCCAT	36730
QY	1756	CAAAAAAAAAAAAAAAAA 1772	
Db	36731	CTCAAAAAAAAAAAAAA 36747	

## RESULT 14

```

US-10-408-168-1/c
Sequence 1, Application US/10408168
Publication No. US20030235847A1
GENERAL INFORMATION:
APPLICANT: Paepfer, Bryan W.
APPLICANT: Proll, Sean
APPLICANT: Charmsley, Patrick R.
APPLICANT: Brunkow, Mary E.
APPLICANT: Ulteerlinden, Andreas Gerardus
TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
FILE REFERENCE: 240083.525
CURRENT APPLICATION NUMBER: US/10/408.168
CURRENT FILING DATE: 2003-04-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 130320
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

```

```

: NAME/KEY: mic feature
: LOCATION: 100722, 100754, 102080, 117731, 124408, 124532, 124585, 124955
: LOCATION: 124956, 124963, 124964, 124965, 124967
: OTHER INFORMATION: n = A,T,C or G
US-10-408-168-1

```

Query Match	5.6%;	Score 124.2;	DB 15;	Length 130320
Best Local Similarity	72.6%;	Pred. No. 3.4e-27;		
Matches 204;	Conservative	0;	Mismatches 68;	Indels 9;
				0

QY	1529	GGGTGTGAGGCAGTGAATATCAACAAGATGACGAGTTTGAACCAACTGGCCAACTATG	1588
Db	127635	GGAGGCCGAGCGGGCGGATCACTCGAAGTCAGAGATTGABACCATCTTGCCCAACATG	127578
QY	1589	ATGAAACCCCATCTCTACCAAAAAAATACAATCAGCTGGCCGTGTGTG---TGC	1643
Db	127575	GTGAAACCCCCGCTCTACTAATAAAAAAACAATAATTAAGCCGGCGGTGTGGCCGACGC	127516
QY	1644	CTGTATGTC--CAACGCAGAGGTTGAGG--GGAAATTGCTTGAACCCAGAGGTGTGTG	1699
Db	127515	CAGTAGCCCACTAATTGGAGGCTTAAGGCAGGAAATTTGCTTGAACCCAGAGGTGTGAG	127456
QY	1700	GTTGCAGTGAATTGAGCACTGCACATCAGCCTGGCGACCGAGTGAATCTGTCTCAA	1759
Db	127455	GTTACAGTAAGATGCGCGCCACTGCACCTCGAGCTGGCGACAGAACGAACTCCGCTCA	127396
QY	1760	AAAAAAAAAAAAAGATGCTCAACTTTTGCCCTCTACT	1800
Db	127395	AAAAACAAAAAACAATAAACCCCAAACTCCCCAGCTATTT	127355

## RESULT 15

```

US-09-764-891-7672/c
; Sequence 7672, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7672
; LENGTH: 14874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7672

```

Query Match	5.6%	Score	123.8	DB	10	Length	14874
Best Local Similarity	78.8%	Pred. No.	1e-27				
Matches 201; Conservative	0	Mismatches	42	Indels	12	Gaps	4

Qy	1529	GGGCTGTGAGGCAAGTAATCA	CAGGAGGT	CAGAGTTTGAGACCAACTGGGCACATG	1588
Db	3589	GGAGGCTGTGAGCGAGCGGATCA	CTTTGAGTCTGAGAGTTTCGAGATTAAGCTTGGCCCAACATG	3530	
Qy	1589	ATGAAACCCCATCTCTACCA	AAAAAATACAAATCAGCTGGCCGTGTGTG--TGCCTG	1646	
Db	3529	GTGAAACCCGTGTCTCTACCA	AAAAAATACAAAAATTAATTGATGATGTGTGTGTGTCATGTCTTG	3470	
Qy	1647	TAGTCCCA----	ACGCAAGAGTTGAG--GGGAGATTGCTTGAACCCAGAGGTGGT	1701	
Db	3469	TAAATCCAGCTACTTGTGGAGCTG	AGGTGGAGAGATCACTTGAACCCAGAGGTGGAGGT	3410	
Qy	1702	TGCAGTG-----AGATTGACCA	ACTGCAATCCAGCCTGGCGAGCGAGTGAAGATGTCTC	1756	
Db	3409	TGCAATGAGCTAAGATTGTGTCC	ATTGTGATCTCCAGCCTGGGCAACAGAGGAGATCTGTTTC	3350	
Qy	1757	AAAAAAAAAAAAAAAA	1771		
Db	3349	AAAAAAAAAAAAAAAA	3335		

Mon Mar 29 09:56:42 2004

us-09-622-964-1\_copy\_13900\_16115.rnpb

Page 8

Search completed: March 25, 2004, 09:52:04  
Job time : 1276 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 02:59:44 ; Search time 5671 Seconds  
(without alignments)  
11668.946 Million cell updates/sec

Title: US-09-622-964-1\_COPY\_13900\_16115

Perfect score: 2216  
Sequence: 1 accctctgtctcttcaccgac.....cctatgagtttaataaac 2216

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estcom:\*  
17: em\_gse\_hum:\*  
18: em\_gse\_inv:\*  
19: em\_gse\_pln:\*  
20: em\_gse\_vic:\*  
21: em\_gse\_fun:\*  
22: em\_gse\_mam:\*  
23: em\_gse\_mus:\*  
24: em\_gse\_pro:\*  
25: em\_gse\_rnd:\*  
26: em\_gse\_pig:\*  
27: em\_gse\_vrl:\*  
28: gb\_gse1:\*  
29: gb\_gse2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	626.2	28.3	669	12	BQ028631 UI-1-EO-
2	619.2	27.9	668	12	CA441116 UI-H-DPO-
3	611.2	27.6	654	13	BU629816 UI-H-FLD-
4	611.2	27.6	669	14	CA308356 UI-H-FTL-

5	603.4	27.2	773	10	BE410951	BE410951 601303662
6	600.6	27.1	629	14	CA397980	CA397980 c98c05.x
7	581.2	26.2	662	12	CG950642	CG950642 MR1-CT073
8	581	26.2	593	10	BE385296	BE385296 601277572
9	555.8	25.1	636	14	CA397981	CA397981 c98c05.y
10	545	24.6	545	14	CA395098	CA395098 c86c05.y
11	530.6	23.9	545	13	BU657324	BU657324 C122H04.z
12	501.8	22.6	711	13	BK095540	BK095540 BX095540
13	496.6	22.4	523	12	BM697639	BM697639 UI-E-DXO-
14	492.6	22.2	529	12	BM707813	BM707813 UI-E-C11-
15	466.4	21.0	469	9	AA307119	AA307119 EST178031
16	457.4	20.6	479	13	BQ185946	BQ185946 UI-E-EJ1-
17	457.4	20.6	484	13	BQ184248	BQ184248 UI-E-EJ1-
18	453.2	20.5	651	13	BU731149	BU731149 UI-E-C11-
19	430.8	19.4	601	12	BG951790	BG951790 MR1-CT073
20	429.8	19.4	466	13	BQ187559	BQ187559 UI-E-EJ1-
21	429	19.4	484	14	N31453	N31453 yx55d08.r1
22	422	19.0	422	9	A1917923	A1917923 t214G04.x
23	414.4	18.7	421	9	A1074304	A1074304 o256b11.x
24	413.2	18.6	444	14	N33227	N33227 y07d02.g1
25	411.6	18.6	494	14	M41943	M41943 y07d02.r1
26	396.2	17.9	766	10	BE275846	BE275846 601121656
27	362.4	16.4	434	12	BI031472	BI031472 PMO-MT043
28	362	16.3	432	12	BI031468	BI031468 PMO-MT043
29	349.2	15.8	486	13	C18742	C18742 C18742 Huma
30	347.4	15.7	474	9	A1190190	A1190190 q437c03.x
31	347.4	15.7	592	12	BM633028	BM633028 UI-E-C10-
32	345.8	15.6	484	13	BU741783	BU741783 UI-E-EJ0-
33	345.8	15.6	659	13	BU730894	BU730894 UI-E-C11-
34	345.4	15.6	537	12	BM691456	BM691456 UI-E-C11-
35	344.2	15.5	674	13	BU731809	BU731809 UI-E-EJ1-
36	342.6	15.5	676	13	BU741926	BU741926 UI-E-EJ0-
37	342.4	15.5	371	12	BM694530	BM694530 UI-E-C11-
38	332.2	15.0	429	12	BI026143	BI026143 CM4-MT036
39	327.2	14.8	599	12	BG951382	BG951382 H2RPE-001
40	325	14.7	419	12	BM718146	BM718146 UI-E-EJ0-
41	310.4	14.0	419	12	BM685122	BM685122 UI-E-EJ1-
42	298.8	13.5	503	12	BM932117	BM932117 UI-E-EJ1-
43	295.4	13.3	657	29	CE720218	CE720218 t19t-g88-
44	295.4	13.3	492	13	BU726009	BU726009 UI-E-C10-

#### ALIGNMENTS

RESULT 1  
LOCUS BQ028631  
DEFINITION UI-1-EO-ayz-f-11-0-UI.g1 NCI CGAP P17 Homo sapiens cDNA clone  
ACCESSION BQ028631  
VERSION BQ028631.1 GI:19763910  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 669)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-r@mail.nih.gov  
Tissue Procurement: ATTC  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, bento-soares@uiowa.edu  
The following repetitive elements were found in this CDNA  
sequence: 393-621, >Alu

Seq primer: M13 FORWARD  
POLYA=Yes

FEATURES  
source

Location/Qualifiers  
1..669  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-1-EE0-ayz-f-11-0-UI"  
/tissue\_type="Placenta Chorioncarcinoma"  
/lab\_host="DH10B (life technologies)"  
/clone\_lib="NCI CGAP P17"  
/note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP P17 is a cDNA library containing the following tissue(s): Placenta Chorioncarcinoma UAR Line (ATCC). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AGGCATCACA.  
TAG\_TISSUE=human chorioncarcinoma  
TAG\_LIB=UI-1-EE0  
TAG\_SEQ=AGGCATCACA"

ORIGIN

Query Match 28.3%; Score 626.2; DB 12; Length 669;  
Best Local Similarity 98.2%; Pred. No. 4.5e-138;  
Matches 642; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

1159 TTGAGGGTGCAGGGCCCTTAGGTGATTTTCTCACTGCTGGGAACCTCAACAAATAC 1218  
17 TTGAGGGTGCAGGGCCCTTAGGTGATTTTCTCACTGCTGGGAACCTCAACAAATAC 1218  
1219 TTCTTCCTTCTTGGGGGTGAGCCCAAGGTGCACAAATATGATATTTCCCTTATTC 1278  
77 TTCTTCCTTCTTGGGGGTGAGCCCAAGGTGCACAAATATGATATTTCCCTTATTC 136  
1279 AGATTCCTGGAACATGTCACCCATTTTAAACCCCACTTACGCCCAATCACTGCG 1338  
137 AGATTCCTGGAACATGTCACCCATTTTAAACCCCACTTACGCCCAATCACTGCG 136  
1339 AGGAAGTGAATCTTCTTCTGATTCACAGAGTACTTTCACGGGTCAAGACAG 1398  
197 AGGAAGTGAATCTTCTTCTGATTCACAGAGTACTTTCACGGGTCAAGACAG 255  
1339 CAGCTATTATGATGAAACCTTAAAGGGCAACATTTTCCTTCTAGGCTAAGA 1458  
256 CAGCTATTATGATGAAACCTTAAAGGGCAACATTTTCCTTCTAGGCTAAGA 315  
1459 CAGGAACCTTGGCAACATCTGTGGCTGTCAGCAAAAGATGTCATATTAAATCTT 1518  
316 CAGGAACCTTGGCAACATCTGTGGCTGTCAGCAAAAGATGTCATATTAAATCTT 375  
1519 GTCTTGGGCTGGGGTGGAGGCAAGTGAATCAGAGAGTTCAGAGTTAGAACCACT 1578  
376 GTCTTGGGCTGGGGTGGAGGCAAGTGAATCAGAGAGTTCAGAGTTAGAACCACT 435  
1579 GGGCAACATGATGAAACCCCATCTTACCAAAAAAATACATGCTGGCCGTCTGG 1638  
436 GGGCAACATGATGAAACCCCATCTTACCAAAAAAATACATGCTGGCCGTCTGG 495  
1639 TGTGCTGTAGTCCCAAGCAGAGAGTGGGGGAATTTCTTGAACCCAGAGGTGT 1698  
496 TGTGCTGTAGTCCCAAGCAGAGAGTGGGGGAATTTCTTGAACCCAGAGGTGT 555  
1699 GGTTCGATGAGATTCAGCAATCCAGCTGAGGGCGAGAGTGAAGCTGTCTCA 1758  
556 GGTTCGATGAGATTCAGCAATCCAGCTGAGGGCGAGAGTGAAGCTGTCTCA 615

QY 1759 AAAAAAAAAAGGATCGTCTCAACCTTGGCTTCACTGCAACATTTGG 1812  
Db 616 AAAAAAAAAAGGATCGTCTCAACCTTGGCTTCACTGCAACATTTGG 669

RESULT 2  
CA441116/c  
LOCUS  
DEFINITION  
UI-H-DP0-ava-c-06-0-UI.61 NCI CGAP P81 Homo sapiens cDNA clone  
CA441116  
VERSION  
KEYWORDS  
CA441116.1 GI:24805536  
EST.  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 668)  
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: Dr. Mary Hendrix  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
The following repetitive elements were found in this cDNA  
sequence: 1-21, >AT-rich#low\_complexity 483-667, >Alu (matched  
complement)  
Seq primer: M13 FORWARD  
POLYA=Yes

JOURNAL  
COMMENT

FEATURES

source

Location/Qualifiers  
1..668  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-DP0-ava-c-06-0-UI"  
/tissue\_type="Fibrosarcoma"  
/lab\_host="DH10B (life technologies)"  
/clone\_lib="NCI CGAP P81"  
/note="Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP P81 is a cDNA library containing the following tissue(s): Fibrosarcoma Cell line HT-1080 (ATCC number CCL-121). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GTTCTACGAG.  
TAG\_TISSUE=fibrosarcoma  
TAG\_LIB=UI-H-DP0  
TAG\_SEQ=GTTCTACGAG"

ORIGIN

Query Match 27.9%; Score 619.2; DB 14; Length 668;  
Best Local Similarity 99.2%; Pred. No. 2.1e-136;  
Matches 633; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1581 CCAACATGATGAAACCCCATCTTACCAAAAAAATAC-ATAATGAGTGGCGGCGTG 1638  
Db 668 CCAACATGATGAAACCCCATCTTACCAAAAAAATACATAATGAGTGGCGGCGTG 609  
QY 1639 TGTGCTGTAGTCCCAAGCAGAGAGTGGGGGAATTTCTTGAACCCAGAGGTGT 1698

```
|||||
Db      608  TGTGCGGTGATGCCCAACGACGAGGAGTTGAGGGGGAATGTCTGAACCCAGGAGGTGGT 549
        |||
Qy      1699 GGTTCAGTGAAGATTGACAACTGCATTCAGCTGGGCGACGAGTGAAGACTGTCTCA 1758
        |||
Db      548  GGTTCAGTGAAGATTGACAACTGCATTCAGCTGGGCGACGAGTGAAGACTGTCTCA 489
        |||
Qy      1759 AAAAAAAAAAAGATCGTCTCAACCTTTGCGCCCTCTACTGCAACATTTTGGTATTTG 1818
        |||
Db      488  AAAAAAAAAAAGATCGTCTCAACCTTTGCGCCCTCTACTGCAACATTTTGGTATTTG 429
        |||
Qy      1819 AAATGAAGTACCTTCCATACCTTAATGCTTTAATATCTTCACTGAAGATGAAGC 1878
        |||
Db      428  AAATGAAGTACCTTCCATACCTTAATGCTTTAATATCTTCACTGAAGATGAAGC 369
        |||
Qy      1879 ACATTCCTAACCTGCTTCTTAATGAGGATGCTTGCAGCCAGGTCTCACTGTGTGTA 1938
        |||
Db      368  ACATTCCTAACCTGCTTCTTAATGAGGATGCTTGCAGCCAGGTCTCACTGTGTGTA 309
        |||
Qy      1939 CACCAGACGACATGATCCAGTCCAGCCATACAGCTGTCACACTGAAGAACGTGTC 1998
        |||
Db      308  CACCAGACGACATGATCCAGTCCAGCCATACAGCTGTCACACTGAAGAACGTGTC 249
        |||
Qy      1999 TACAACAGCCTGATCAATATGTTAGCTTAATATGATTAATCCAGACTACTTCAAGCCT 2058
        |||
Db      248  TACAACAGCCTGATCAATATGTTAGCTTAATATGATTAATCCAGACTACTTCAAGCCT 189
        |||
Qy      2059 TTATATGCTTTTATTTCTAATAAACTGTGAAGCTGACAGTCACTTGAACATTTTAC 2118
        |||
Db      188  TTATATGCTTTTATTTCTAATAAACTGTGAAGCTGACAGTCACTTGAACATTTTAC 129
        |||
Qy      2119 TCAGACTCTGATTCAGAGTCGAGGAACCCCTTAGTCTATGTAATCCAGACGACAC 2178
        |||
Db      128  TCAGACTCTGATTCAGAGTCGAGGAACCCCTTAGTCTATGTAATCCAGACGACAC 69
        |||
Qy      2179 CTTAGTATACTGCCCCAACTAATGAGTTTAAATTAATAC 2216
        |||
Db      68  CTTAGTATACTGCCCCAACTAATGAGTTTAAATTAATAC 31
        |||

RESULT 3
BU629816/c 654 bp mRNA linear EST 23-SEP-2002
LOCUS      BU629816
DEFINITION UI-H-FLO-bdm-e-06-0-UI-61 NCI CGAP FLO Homo sapiens cDNA clone
            UI-H-FLO-bdm-e-06-0-UI 3', mRNA sequence.
ACCESSION  BU629816
VERSION    BU629816.1 GI:23296239
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 654)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: James Martin
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Clone distribution information can be obtained
            from Dr. M. Bento Soares, bento-soares@uiowa.edu
            The following repetitive elements were found in this cDNA
            sequence: 303-531. >ALU (matched complement)
            Seq primer: M13 FORWARD
            PolyA=yes.

FEATURES
            source
            1..654
            /organism="Homo sapiens"
            /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="UI-H-FLO-bdm-e-06-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="NCI CGAP FLO"
/clone_lib="NCI CGAP FLO"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
a pool of mRNA obtained from 4 cell lines from grade III
chondrosarcoma tissues. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tag for this library is GAGGTGCGTG. The cell line
was provided by Dr James Martin from University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_lib=UI-H-FLO
TAG_SEQ=GAGGTGCGTG"

ORIGIN
Query Match 27.6%; Score 611.2; DB 13; Length 654;
Best Local Similarity 98.4%; Prid. No. 1.7e-134;
Matches 627; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy      1414 AAACCTTAAAGGGGCAACATTTCAATCTTCTTCTAGGCTAGACGAACTTGGCAA 1473
        |||
Db      654  AAACCTTAAAGGGGCAACATTTCAATCTTCTTCTAGGCTAGACGAACTTGGCAA 595
        |||
Qy      1474 CATCTGTGCGCTGTTGAGCA-A-AGATGTTCAATATTTAAGATCTTGTCTTGGCTGGGT 1532
        |||
Db      594  CATCTGTGCGCTGTTGAGCAANAGATTTAATTTAAGATCTTGTCTTGGCTGGGT 535
        |||
Qy      1533 GTGAGAGGCAAGTGAATCAACGAGAGTCAGAGATTGAGACCACTGGCCAACTGATGA 1592
        |||
Db      534  GTGAGAGGCAAGTGAATCAACGAGAGTCAGAGATTGAGACCACTGGCCAACTGATGA 475
        |||
Qy      1593 AACCCCATCTCACAAAAAATAAATCAATCAGCTGCGGTGCTGCTGCTGATGTC 1652
        |||
Db      474  AACCCCATCTCACAAAAAATAAATCAATCAGCTGCGGTGCTGCTGCTGATGTC 415
        |||
Qy      1653 CAACGACGAGGTTGAGGGGAGAAATGCTTGAACCCAGAGAGTGTGTTCAGTAGAT 1712
        |||
Db      414  CAACGACGAGGTTGAGGGGAGAAATGCTTGAACCCAGAGAGTGTGTTCAGTAGAT 355
        |||
Qy      1713 TGAGCACTGCAATTCAGCTTGGGCGACGAGTGAAGTGTCTCAAAAAAAAAAAAA 1772
        |||
Db      354  TGAGCACTGCAATTCAGCTTGGGCGACGAGTGAAGTGTCTCAAAAAAAAAAAAA 295
        |||
Qy      1773 GGAATGCTCAACCTTTGCGCCCTCTACTGCAACATTTTGGTATTTGAATGAAGTACCT 1832
        |||
Db      294  GGAATGCTCAACCTTTGCGCCCTCTACTGCAACATTTTGGTATTTGAATGAAGTACCT 235
        |||
Qy      1833 TCCATATCTTATGCTGTTAATATCTTCAATCTGACTAGGAGTGAAGACATTCCTAACCTG 1892
        |||
Db      234  TCCATATCTTATGCTGTTAATATCTTCAATCTGACTAGGAGTGAAGACATTCCTAACCTG 175
        |||
Qy      1893 CTTCTTAATGAGGATGCTTGGCAGCAGAGTCTCACTGTGTGTAACCAAGACAGAC 1952
        |||
Db      174  CTTCTTAATGAGGATGCTTGGCAGCAGAGTCTCACTGTGTGTAACCAAGACAGAC 115
        |||
Qy      1953 TGATCAGACGACGATCAGCTGTCCACACTGAAGAGGTCTTCAACAGCTGAA 2012
        |||
Db      114  TGATCAGACGACGATCAGCTGTCCACACTGAAGAGGTCTTCAACAGCTGAA 55
        |||
Qy      2013 TCAAAATGTTAGCTTAATGATTAATCCAGACTA 2049
        |||
Db      54  TCAAAATGTTAGCTTAATGATTAATCCAGACTA 18
        |||
```

RESULT 4  
CA308356/c  
LOCUS  
DEFINITION CA308356 669 bp mRNA linear EST 01-NOV-2002  
UI-H-PT1-bhz-k-14-0-UI.s1 NCI CGAP PT1 Homo sapiens cDNA clone  
CA308356  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 669)  
NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapdb-remail.nih.gov](mailto:cgapdb-remail.nih.gov)  
Tissue Procurement: Dr. Gary W. Hunninghake, U of I  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
The following repetitive elements were found in this cDNA  
sequence: 481-669, >ALU (matched complement)  
Seq primer: M13 FORWARD  
PolyA=yes.

FEATURES  
source  
1..669  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-PT1-bhz-k-14-0-UI"  
/tissue\_type="Aveolar Macrophage"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
NCI CGAP PT1 is a normalized cDNA library constructed from  
a pool of 81 RNA samples from Alveolar Macrophages  
challenged with different treatments. The library was  
normalized according to Bonaldo, Lennon and Soares, Genome  
Research, 6:791-806, 1996. First strand cDNA synthesis was  
primed with an oligo-dT primer containing a Not I site.  
Double stranded cDNA was ligated to an EcoR I adaptor,  
pT7T3-Pac vector. The oligonucleotide used to prime the  
synthesis of first-strand cDNA contains a library tag  
(GATC18 tail). The sequence tag for this library is  
GGCCATGCCG. The tissue was provided by Dr. Gary W.  
Hunninghake of the University of Iowa.  
TAG TISSUE=Human Lung Aveolar Macrophage  
TAG LIB=UI-H-PT1  
TAG\_SEQ=GGCCATGCCG"

ORIGIN  
Query Match 27.6%; Score 611.2; DB 14; Length 669;  
Best Local Similarity 99.1%; Pred. No. 1.7e-134;  
Matches 635; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 1577 CTGGCCACATGATGAAACCCCATCTTA-CCAAAAAATATCAATCAGCTGCGCTCG 1635  
DB 669 CTGGCCACATGATGAAACCCCATCTTA-CCAAAAAATATCAATCAGCTGCGCTCG 610  
QY 1636 TGGTGGCTCTGAGTCCCAAGCAGAGTGAAGGAGGAGTCTTGAACCAAGAGGT 1695  
DB 609 TGGTGGCTCTGAGTCCCAAGCAGAGTGAAGGAGGAGTCTTGAACCAAGAGGT 550

QY 1696 GGTGTTGAGTGAGATTGAGCACTGCAATCCAGCTGGGCGACGAGTGAAGCTGTCT 1755  
DB 549 GGTGTTGAGTGAGATTGAGCACTGCAATCCAGCTGGGCGACGAGTGAAGCTGTCT 490  
QY 1756 CAAAAAAGAAAAAAGAAATGCTGCACTTGGCTGCTGCACTGCACTTGGTAT 1815  
DB 489 C-AAAAAAAAAAAAAAAAAGAAATGCTGCACTTGGCTGCTGCACTGCACTTGGTAT 431  
QY 1816 TTGAATGAGAGTACCTTCCTGCACTTGGCTGCTGCACTTGGTAT 1875  
DB 430 TTGAATGAGAGTACCTTCCTGCACTTGGCTGCTGCACTTGGTAT 371  
QY 1876 AGCATTCTTAACTGCTTCCCAATGAGGAGTCTTGGCTGCTGCACTTGGTAT 1935  
DB 370 AGCATTCTTAACTGCTTCCCAATGAGGAGTCTTGGCTGCTGCACTTGGTAT 311  
QY 1936 GTACACAGAGAGACAGTATGCTGCACTGCTGCTGCACTGCTGCACTGCTGCTGCT 251  
DB 310 GTACACAGAGAGACAGTATGCTGCACTGCTGCTGCACTGCTGCACTGCTGCTGCT 215  
QY 1996 TCTTCAACAGCTGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2055  
DB 250 TCTTCAACAGCTGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191  
QY 2056 CCTTAAAGCTCTTAACTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2115  
DB 190 CCTTAAAGCTCTTAACTGCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131  
QY 2116 AACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2175  
DB 130 AACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71  
QY 2176 CACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2216  
DB 70 CACCTTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 30

RESULT 5  
LOCUS  
DEFINITION BE410951 773 bp mRNA linear EST 21-JUN-2000  
601303662F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3638175 5',  
mRNA sequence.  
ACCESSION  
VERSION BE410951  
KEYWORDS  
SOURCE BE410951.1 GI:9347401  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 773)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1997)  
JOURNAL  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapdb-remail.nih.gov](mailto:cgapdb-remail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNM at: [image.llnl.gov](mailto:image.llnl.gov)  
Plate: LHCW38 row: 1 column: 16  
High quality sequence stop: 662.

FEATURES  
source  
1..773  
/location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3638175"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_21"



/note="Organ: placentae; Vector: pOTB1; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGGAG(G). Size-selected for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

```
Query Match      27.2%; Score 603.4; DB 10; Length 773;
Best Local Similarity 96.9%; Pred. No. 1.2e-132;
Matches 648; Conservative 0; Mismatches 16; Indels 5; Gaps 3;

QY 68 GGAATGCTCAGGCTGGGCTTCTTGGCCCTTCTTGAAGCTTGCATGCTCCATATCATCATCC 127
    |||||
DB 1 GGAATGCTCAGGCTAGCATCTTGGCCCTTCTTGAAGCTTGCATGCTCCATATCATCATCC 60

QY 128 TCCGAGGCGAATCTCAAGACCAAACTAGTGGCCCAAGAGGGGAATCCCTTCCACGA 187
    |||||
DB 61 TCCGAGGCGAATCTCAAGACCAAACTAGTGGCCCAAGAGGGGAATCCCTTCCACGA 120

QY 188 GGGCCTGCCC-AAAAACACAGGACGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 246
    |||||
DB 121 GGGCCTGCCCACCAAAACACAGGACGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180

QY 247 AGGCTGTGAAGCTTAAAGCTGTGAGCCCTTCAAGTGTGCCCCAATGTATCAGAGGCGAG 306
    |||||
DB 181 AGGCTGTGAAGCTTAAAGCTGTGAGCCCTTCAAGTGTGCCCCAATGTATCAGAGGCGAG 240

QY 307 GCTACTACAGTGGCCCCACAGAGGCGCCCTGAGCCCACTCCCATGTTCTTCCCTTAGAAC 366
    |||||
DB 241 GCTACTACAGTGGCCCCACAGAGGCGCCCTGAGCCCACTCCCATGTTCTTCCCTTAGAAC 300

QY 367 CATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGAGATAGACACCAAGAGAGAGAGAGAGAG 426
    |||||
DB 301 CATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGAGATAGACACCAAGAGAGAGAGAGAGAG 360

QY 427 AGACTGTGAGTTCGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAG 486
    |||||
DB 361 AGACTGTGAGTTCGGGGCCCAAGAAAGTTTGAATGCTCTCAGAGAGAGAGAGAGAGAGAG 420

QY 487 TGAATGAGACCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
    |||||
DB 421 TGAATGAGACCAAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

QY 547 ATATGCGAGAGATCCCGGAAATATCACTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
    |||||
DB 481 ATATGCGAGAGATCCCGGAAATATCACTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 607 TACACACTACATCAAGATCAATGAGATCTTATGAGGCTTGGAAACAGAGTCTGTCC 666
    |||||
DB 541 TACACACTACATCAAGATCAATGAGATCTTATGAGGCTTGGAAACAGAGTCTGTCC 600

QY 667 TCACACTGAACCAAGGGGCACTGATGCGCTGTGCCCCACCCAGCTTCCCTTGTCTGA 726
    |||||
DB 601 TCACACTGAACCA-GGGCACTGATGCGC--TGTGCCCCAACCCAGCTTCTTGTCTGA 656

QY 727 GCCTACCTT 735
    |||||
DB 657 GCCTACCTT 665
```

RESULT 6  
CA397980/c 629 bp mRNA linear EST 06-NOV-2002  
LOCUS CA397980.1 Human Retinal pigment epithelium/choroid cDNA  
DEFINITION (Un-normalized, unambiguated): c8 Homo sapiens cDNA clone c898c05

ACCESSION CA397980  
VERSION CA397980.1 GI:24735787  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

## REFERENCE

1 (bases 1 to 629)

## AUTHORS

Wistow,G., Bernstein,S.T., Wyatt,M.K., Farris,R.N., Behal,A.,

## TITLE

Touchman,G.W., Bouffard,G., Smith,D., and Peterson,K.

Expressed sequence tag analysis of human RPE/choroid for the

NIH Bank Project: Over 6000 non-redundant transcripts, novel genes

and splice variants

## JOURNAL

Mol. Vis. 8 (4), 205-220 (2002)

## MEDLINE

22103460

## PubMed

12107410

## COMMENT

Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: graeme@helix.nih.gov  
Plate: 98 row: c column: 05  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers

## FEATURES

## source

1..629

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="c898c05"

/tissue\_type="RPE/choroid"

/dev\_stage="Adult"

/lab\_host="EMD108"

/clone\_lib="Human Retinal pigment epithelium/choroid cDNA

(un-normalized, unambiguated): c8"

/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor

eyes (75-80 years old) yielded approximately 600 mg of

dissected RPE/choroid tissue. This in turn yielded 340 ug

of total RNA and 7 ug of mRNA. A directionally cloned cDNA

library in the pCMVSPORT6 vector was constructed at Life

Technologies (Rockville, MD; now part of Invitrogen Corp),

essentially following the protocols of the SuperScript

Plasmid System (Invitrogen Corp).

&lt;http://www.invitrogen.com/&gt;. The library code

designation was c8. For this library, cDNA inserts were

cloned into the NotI/Mlu sites of the vector. EST

analysis was performed on the unambiguated library at the

NIH Intramural Sequencing Center (NISC)."

## ORIGIN

```
Query Match      27.1%; Score 600.6; DB 14; Length 629;
Best Local Similarity 98.2%; Pred. No. 5.6e-132;
Matches 617; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 994 GCCCAACTTACTTTGAGCAAGGGTGGCTGACCAAAACCATGAGGTGAGTCACTGG 1053
    |||||
DB 629 GCCCAACTTACTTTGAGCAAGGGTGGCTGACCAAAACCATGAGGTGAGTCACTGG 570

QY 1054 ATGACGATGAAACATTTCCCTTACTTATTTAGGTATGATCCCAAGCACTAAGAGAAAG 1113
    |||||
DB 569 ATGACGATGAAACATTTCCCTTACTTATTTAGGTATGATCCCAAGCACTAAGAGAAAG 510

QY 1114 GGTGAGAGAACTGCTCACTCCTTAGAAGCTGGTAGAGTGGAGTGGAGGCTGACG 1173
    |||||
DB 509 GGTGAGAGAACTGCTCACTCCTTAGAAGCTGGTAGAGTGGAGTGGAGGCTGACG 450

QY 1174 GCCCTTAGTCAATTTTCTCAGTGGCTGGAAGCTGACCAAAATATCTTCTTCTTGG 1233
    |||||
DB 449 GCCCTTAGTCAATTTTCTCAGTGGCTGGAAGCTGACCAAAATATCTTCTTCTTGG 390

QY 1234 GGTGAGCCCAAGCTGTCAAAATATGATATTTCCCTTATTCAGATTTCTTGGACAC 1293
    |||||
DB 389 GGTGAGCCCAAGCTGTCAAAATATGATATTTCCCTTATTCAGATTTCTTGGACAC 330

QY 1294 TGTCAACCAATTAACAACCCCACTTCAAGCCCAATCAAGTGGAGAGTAACTTC 1353
    |||||
```

Db 329 TTTCACCAATATTAACACCCCACTTACGCCCAATCAGTGGAGAGTGTACTTC 270  
QY 1354 CCTTTTGTGATTCAGAGCTTATCTTTCACGGGTGCAAGACAGAGCATTTATGATG 1413  
Db 269 CCTTTTGTGATTCAGAGCTTATCTTTCACGGGTGCAAGACAGAGCATTTATGATG 211  
QY 1414 AACCTTAAAGGCAACATTTGATCTTTCAGGTTAAGACAGAACTTGGCAA 1473  
Db 210 AAACTTAAAGGCAACATTTGATCTTTCAGGTTAAGACAGAACTTGGCAA 151  
QY 1474 CATCTGTGCTCTTTCAGGAAAGATTTCAATTTAAGATCTTTCAGGCTGGGTG 1533  
Db 150 CATCTGTGCTCTTTCAGGAAAGATTTAATTTAAGATCTTTCAGGCTGGGTG 91  
QY 1534 TGGAGCAAGTGAATCAGAGAGTTCAGGTTTGAACCACTGGGCAATGATGAA 1593  
Db 90 TGGAGCAAGTGAATCAGAGAGTTCAGGTTTGAACCACTGGGCAATGATGAA 31  
QY 1594 ACCCACTCTCTACCAAAAAATATGAAA 1621  
Db 30 ACCCACTCTCTACCAAAAAATATGAAA 3

RESULT 7  
Bg950642 662 bp mRNA linear EST 12-JUN-2001  
LOCUS Bg950642  
DEFINITION MRL-CT0735-080101-001-408 CT0735 Homo sapiens cDNA, mRNA sequence.  
ACCESSION Bg950642  
VERSION Bg950642.1 GI:14368813  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Gouldman, G.H., de Oliveira, A.F., Matsukuma, A., Bais, G.S., Simpson, D.H.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PubMed 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MRLct2-MRL-CT0735-  
080101-001-408&t3=2001-01-08&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 593.  
Location/Qualifiers  
1..662

FEATURES  
Source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="CT0735"  
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the puc 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low

## ORIGIN stringency conditions."

Query Match 26.2%; Score 581.2; DB 12; Length 662;  
Best Local Similarity 96.8%; Pred. No. 2,3e-127;  
Matches 603; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 1152 GGTGAGTTGAGGGGTGTCAGGCGCCCTTACGTATTTTTCATGCTGGGAACCTGACC 1211  
Db 11 GGTGAGTTGAGGGGTGTCAGGCGCCCTTACGTATTTTTCATGCTGGGAACCTGACC 70  
QY 1212 AAAATCTCTCTGCTTCTTGGGAGTACGCCCAAGCTGTCACAAATATGATTTCCCT 1271  
Db 71 AAAATCTCTGCTTCTTGGGAGTACGCCCAAGCTGTCACAAATATGATTTCCCT 130  
QY 1272 TTATTCAGATTTCTGACACATGTAACCAATTTTAAACCCCACTTCAGGCCCAATC 1331  
Db 131 TTATTCAGATTTCTGACACATGTAACCAATTTTAAACCCCACTTCAGGCCCAATC 190  
QY 1332 ACCGAGGAGGAGTGTAACTTCTCTTCTGATTTCTCAGATTTCTTCAAGGCTCA 1391  
Db 191 ACCGAGGAGGAGTGTAACTTCTCTTCTGATTTCTCAGATTTCTTCAAGGCTCA 249  
QY 1392 GAACAGCAGCTATTTATGTAAGAACTTAAAGGCAATTTTCAATCTTCTTAG 1451  
Db 250 GAACAGCAGCTATTTATGTAAGAACTTAAAGGCAATTTTCAATCTTCTTAG 309  
QY 1452 GCTAAGACAGAACTTGGCAACATCTGTGGCTGTTCAGCAAGAGATGTTCAATTTAA 1511  
Db 310 GCTAAGACAGAACTTGGCAACATCTGTGGCTGTTCAGCAAGAGATGTTCAATTTAA 369  
QY 1512 GAATTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGGTTCAGAGTTTGA 1571  
Db 370 GAATTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGGTTCAGAGTTTGA 429  
QY 1572 CCAACCTGGCCCAATGATTAACCCCATCTTACCAAAAAATATACAACTGAGTGGCC 1631  
Db 430 CCAACCTGGCCCAATGATTAACCCCATCTTACCAAAAAATATACAACTGAGTGGCC 489  
QY 1632 GTCTGTGTGCTGCTGTGATCCCAAGCAGAGGTTGAGGGAGAAATGCTTGAACCCAG 1691  
Db 490 GTCTGTGTGCTGCTGTGATCCCAAGCAGAGGTTGAGGGAGAAATGCTTGAACCCAG 549  
QY 1692 AGGTGTGTGTTGAGAGATTTGACCACTGCAATTCAGCTGGGCGACGAGTGAAGACT 1751  
Db 550 AGGTGTGTGTTGAGAGATTTGACCACTGCAATTCAGCTGGGCGACGAGTGAAGACT 609  
QY 1752 GTTCAAAAAAATTTTAAAGG 1774  
Db 610 GTTCAAAAAAATTTTAAAGG 632

RESULT 8  
BE385296 593 bp mRNA linear EST 21-JUL-2000  
LOCUS BE385296  
DEFINITION MRL-27757221 NIH-MGC 20 Homo sapiens cDNA clone IMAGE361865 5',  
mRNA sequence.  
ACCESSION BE385296  
VERSION BE385296.1 GI:9330661  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
TITLE NIH-MGC http://mgi.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgaabs@mail.nih.gov  
Tissue Procurement: ATCC/CDTD/DRP  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov  
plate: LLCM287 row: 0 column: 18  
High quality sequence stop: 593.

## FEATURES

## Source

Location/Qualifiers

1. 593  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:361865"  
/tissue\_type="melanotic melanoma"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NIH MGC 20"  
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 26.2%; Score 581; DB 10; Length 553;  
Best Local Similarity 99.8%; Pred. No. 2,6e-127;  
Matches 592; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

## FEATURES

## Source

Location/Qualifiers

1. 636  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="c998c05"  
/tissue\_type="RPE/choroid"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/clone\_1lb="Human Retinal pigment epithelium/choroid cDNA  
(un-normalized, unamplified): cs"  
/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor  
eyes (75-80 years old) yielded approximately 600 mg of  
dissected RPE/choroid tissue. This in turn yielded 340 ug  
of total RNA and 7 ug of mRNA. A directionally cloned cDNA  
library in the pCMVSPORT6 vector was constructed at Life  
Technologies (Rockville, MD; now part of Invitrogen Corp),  
essentially following the protocols of the Superscript  
Plasmid System (Invitrogen Corp).  
<http://www.invitrogen.com/>. The library code  
designation was cs. For this library, cDNA inserts were  
cloned into the NotI/MluI sites of the vector. EST  
analysis was performed on the unamplified library at the  
NIH Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 25.1%; Score 555.8; DB 14; Length 636;  
Best Local Similarity 99.6%; Pred. No. 2,6e-121;  
Matches 557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 9  
CA397981 636 bp mRNA linear EST 06-NOV-2002  
LOCUS CA397981  
DEFINITION c998c05.y1 Human Retinal pigment epithelium/choroid cDNA

QY 17 CAGCTGAACAAAGAGAGAGAGTTCCAGCCCAATCAGAGAGAGAGAGAGAGTCTCA 76  
DB 78 CAGCTGAACAAAGAGAGAGAGTTCCAGCCCAATCAGAGAGAGAGAGAGAGTCTCA 137  
QY 77 CGGTGCATCATTTGGCGGCTTCTTCTAGGCTGAGTCCCATGATCAACATCTCTCCAGGGC 136  
DB 138 CGGTGCATCATTTGGCGGCTTCTTCTAGGCTGAGTCCCATGATCAACATCTCTCCAGGGC 197  
QY 137 AAAGTCAAGAGCAAACTACTGTGGCCCAAGAGAGAGTCTTCTTCAAGAGAGAGTCTG 196  
DB 198 AAAGTCAAGAGCAAACTACTGTGGCCCAAGAGAGAGTCTTCTTCAAGAGAGAGTCTG 257  
QY 197 CAAAAACCAAG 256  
DB 258 CAAAAACCAAG 317

QY 257 GCTTAAGGCTGTGAGCGCTTCAAGTCTTCCCTCACTGTATCATGAGGCCAGGCTACTACAG 316  
 DB 318 GTTAAAGGCTGTGAGCGCTTCAAGTCTTCCCTCACTGTATCATGAGGCCAGGCTACTACAG 377  
 QY 317 TGCCCAAGAGAGCGCCCTCAGCGCCCACTCCCATGTTTCTTCCCTCAGAAACCATGAGGCC 376  
 DB 378 TGCCCAAGAGAGCGCCCTCAGCGCCCACTCCCATGTTTCTTCCCTCAGAAACCATGAGGCC 437  
 QY 377 GTCAAGGCTTCACTGTATCATGAGCGCATATGACACCAAGACAAAGCTTAAAGCTGTAG 436  
 DB 438 GTCAAGGCTTCACTGTATCATGAGCGCATATGACACCAAGACAAAGCTTAAAGCTGTAG 497  
 QY 437 TTCTGGGGCCCAAGAAAGTTTGAATGTCTCTCAGAGAGCGATGGGGCTTGAATGAGCA 496  
 DB 498 TTCTGGGGCCCAAGAAAGTTTGAATGTCTCTCAGAGAGCGATGGGGCTTGAATGAGCA 557  
 QY 497 CCCAGAGATATCTCAAGTGTAGAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 556  
 DB 558 CCCAGAGATATCTCAAGTGTAGAGAGAAAGCTGTGAGTTTAACTGACGATATGCCAGA 617  
 QY 557 GATCCCGGAAATACCTTC 575  
 DB 618 GATCCCGGAAATACCTTC 636

RESULT 10  
 CA395098 545 bp mRNA linear EST 06-NOV-2002  
 LOCUS CA395098  
 DEFINITION cs60c05.Y1 Human Retinal pigment epithelium/choroid cDNA  
 (Un-normalized, unamplified): cs Homo sapiens cDNA clone cs60c05  
 5', mRNA sequence.  
 CA395098  
 EST. CA395098.1 GI:24730259

ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 545)  
 Wiatow, G., Bernstein, S.L., Wiat, M.K., Parria, R.N., Behal, A.,  
 Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.  
 Expressed sequence tag analysis of human RPE/choroid for the  
 NEIBank Project: Over 6000 non-redundant transcripts, novel genes  
 and splice variants  
 M01. Vis. 8 (4), 205-220 (2002)  
 22103460  
 12107410  
 COMMENT Contact: Wiatow G  
 Section on Molecular Structure and Function  
 National Eye Institute  
 6/331, NIH, Bethesda, MD 20892-2740, USA  
 Tel: 301 402 3452  
 Fax: 301 496 0078  
 Email: gtraeme@helix.nih.gov  
 Plate: 60 row: c column: 05  
 Seq primer: M13RPI reverse primer (AB1).  
 Location/Qualifiers  
 1. 545

FEATURES  
 source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="cs60c05"  
 /tissue\_type="RPE/choroid"  
 /dev\_stage="Adult"  
 /lab\_host="EMDH1.08"  
 /clone\_lib="Human Retinal pigment epithelium/choroid cDNA  
 (Un-normalized, unamplified): cs  
 /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor  
 eyes (75-80 years old) yielded approximately 600 mg of  
 dissected RPE/choroid tissue. This in turn yielded 340 ug  
 of total RNA and 7 ug of mRNA. A directionally cloned cDNA  
 library in the pCMVSPORT6 vector was constructed at Life  
 Technologies (Rockville, MD; now part of Invitrogen Corp),  
 Technologies (Rockville, MD; now part of Invitrogen Corp),

ORIGIN  
 Query Match 24.6%; Score 545; DB 14; Length 545;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-119;  
 Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 essentially following the protocols of the SuperScript  
 Plasmid System (Invitrogen Corp).  
 <http://www.invitrogen.com/>. The library code  
 designation was cs. For this library, cDNA inserts were  
 cloned into the NotI/MluI sites of the vector. EST  
 analysis was performed on the unamplified library at the  
 NIH Intramural Sequencing Center (NISC)."

QY 167 GAGGGAATCCCTTCTTCACAGAGGCTGCCCCAAAACCAAGAGCCCAAGACGT 226  
 DB 1 GAGGGAATCCCTTCTTCACAGAGGCTGCCCCAAAACCAAGAGCCCAAGACGT 60  
 QY 227 TAGGGGCGAGGAAGAACAGAGCCCTGGAAGCTTAAAGCTGTGAGAGCGCTTCAAGCTGC 286  
 DB 61 TAGGGGCGAGGAAGAACAGAGCCCTGGAAGCTTAAAGCTGTGAGAGCGCTTCAAGCTGC 120  
 QY 287 CCCACTGTATCAGAGGCGCAGGCTACTACAGTGCCTCCACAGAGCCCTCAGGCCCACTCC 346  
 DB 121 CCCACTGTATCAGAGGCGCAGGCTACTACAGTGCCTCCACAGAGCCCTCAGGCCCACTCC 180  
 QY 347 CATGTTCTTCCCTCTGAACCATACAGCCCGTCAAAAGCTTCAAGTGTACAGCATAGA 406  
 DB 181 CATGTTCTTCCCTCTGAACCATACAGCCCGTCAAAAGCTTCAAGTGTACAGCATAGA 240  
 QY 407 CACCAAGACAAAGAGCTTAAAGAGTGTAGTCTGAGGGCGAAGAAAGTTTGAATGCT 466  
 DB 241 CACCAAGACAAAGAGCTTAAAGAGTGTAGTCTGAGGGCGAAGAAAGTTTGAATGCT 300  
 QY 467 CTCAGAGCGATGAGGCGCTTGAATGAGAGCACCAGAAATATCTCAAGTGTAGAGAGAAAC 526  
 DB 301 CTCAGAGCGATGAGGCGCTTGAATGAGAGCACCAGAAATATCTCAAGTGTAGAGAGAAAC 360  
 QY 527 TGTGAGAGTTAACTGAGAGGATATGAGGATATGAGGATATGAGGATATGAGGATATGAGG 586  
 DB 361 TGTGAGAGTTAACTGAGAGGATATGAGGATATGAGGATATGAGGATATGAGGATATGAGG 420  
 QY 587 GGAACATATCACCAACCAATACATACATACATCACTCAAGATGATGATGATGATGATGATG 646  
 DB 421 GGAACATATCACCAACCAATACATACATACATCACTCAAGATGATGATGATGATGATGATG 480  
 QY 647 CTTGAAAAACAGGTTGTCTCTCACCTGAACCGAGGCGCATGATGCTGCTGCTGCTGCTG 706  
 DB 481 CTTGAAAAACAGGTTGTCTCTCACCTGAACCGAGGCGCATGATGCTGCTGCTGCTGCTG 540  
 QY 707 CCCAG 711  
 DB 541 CCCAG 545

RESULT 11  
 BU657324 545 bp mRNA linear EST 30-SEP-2002  
 LOCUS BU657324  
 DEFINITION cl22h04.Z1 Hembase, Erythroid Precursor Cells (LCB:cl library) Homo  
 sapiens cDNA clone cl22h04 5', mRNA sequence.  
 BU657324  
 VERSION BU657324.1 GI:23369506  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 545)  
 Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.  
 Gene Expression in Human Erythroid Precursor Cells  
 Unpublished (2002)  
 CONTACT: Jeffery L. Miller  
 Laboratory of Chemical Biology  
 National Institute of Diabetes and Digestive and Kidney Diseases



Best Local Similarity 99.4%; Pred. No. 1.9e-108;  
Matches 514; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 148 CCAACTACTGTGGCCCAAG-AGGGAATCCCTTCTCCAGAGAGGCTGCCCAAAACAC 206  
Db 1 CCAACTACTGTGGCCCAAGAGAGGATCCCTTCTCCAGAGAGGCTGCCCAAAACAC 60

QY 207 AAGGCAAGCCAAACAGACGTTAGGGGCGCAGAGAGCAACAGAGCTTGAAGCTTAAGGCT 266  
Db 61 AAGGCAAGCCAAACAGACGTTAGGGGCGCAGAGAGCAACAGAGCTTGAAGCTTAAGGCT 120

QY 267 GTGAGAGCTTCAAGTCTGCCCACTGTATTCAGAGGCGCAGGCTACTTACAGTGCACAG 326  
Db 121 GTGAGAGCTTCAAGTCTGCCCACTGTATTCAGAGGCGCAGGCTACTTACAGTGCACAG 180

QY 327 AGCCCTCAGGCCCCCAGCTCCCATGTTCTTCCCTTGAAGACATCAGCCCGCTCAAGCTT 386  
Db 181 AGCCCTCAGGCCCCCAGCTCCCATGTTCTTCCCTTGAAGACATCAGCCCGCTCAAGCTT 240

QY 387 CACAGTGTACAGGCAATAGACCAAGAGCAAAAGCTTAAAGCTGTAGTTCTGGGGCC 446  
Db 241 CACAGTGTACAGGCAATAGACCAAGAGCAAAAGCTTAAAGCTGTAGTTCTGGGGCC 300

QY 447 AAGAAAGTTTGAATGCTCTCAGAGAGGCAATGGGCGCTTGTAGAGACCCAGAAATA 506  
Db 301 AAGAAAGTTTGAATGCTCTCAGAGAGGCAATGGGCGCTTGTAGAGACCCAGAAATA 360

QY 507 TCTCAAGTAGAGGAGAAAGCTGTGAGTTTAACTGACGATATGCGCAGAGATCCCGAA 566  
Db 361 TCTCAAGTAGAGGAGAAAGCTGTGAGTTTAACTGACGATATGCGCAGAGATCCCGAA 420

QY 567 AATCACTCTCAAGAACTTTGGAAACATACCAACCAACATACACACTACCTCAAGAT 626  
Db 421 AATCACTCTCAAGAACTTTGGAAACATACCAACCAACATACACACTACCTCAAGAT 480

QY 627 CACATGATCCTTATTTGGGCGCTTGGAAACAGGCTGTG 663  
Db 481 CACATGATCCTTATTTGGGCGCTTGGAAACAGGATG 517

RESULT 13  
LOCUS BM697639 523 bp mRNA linear EST 28-FEB-2002  
DEFINITION UI-E-DX0-agn-n-07-0-UI r1 UI-E-DX0 Homo sapiens cDNA clone  
ACCESSION BM697639  
VERSION BM697639.1 GI:19010897  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 523)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL MEDLINE  
PUBMED 97044477  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MBRRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA

FEATURES  
source  
sequence: 9-237, >ALU  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1..523  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-DX0-agn-n-07-0-UI"  
/tissue\_type="fetal eyes"  
/dev\_stage="fetal"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
modified polylinker; Site: 1: EcoR I; Site: 2: Not I;  
UI-E-DX0 is a cDNA library containing the following  
tissue(s): fetal eyes. The library was constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoR I adaptor, digested  
with Not I, and cloned directionally into pT73-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is AGAATCAAGA. This library  
was created for the program, Gene Discovery in the Visual  
System, supported by National Eye Institute (NEI)."

Query Match 22.4%; Score 496.6; DB 12; Length 523;  
Best Local Similarity 98.8%; Pred. No. 3.2e-107;  
Matches 510; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

## ORIGIN

QY 1536 GAGGCAAGTGAATCAAGAGAGTCAAGAGTTTGAAGACCAACCTGGCCAATGATGAAC 1595  
Db 9 GAGGCAAGTGAATCAAGAGAGTCAAGAGTTTGAAGACCAACCTGGCCAATGATGAAC 68

QY 1596 CCCATCTCAACAAAAAATATCAAAATGAGTGGCGCGCGTGTGTGCTTATCTCCAA 1655  
Db 69 CCCATCTCAACAAAAAATATCAAAATGAGTGGCGCGCGTGTGTGCTTATCTCCAA 128

QY 1656 GCGAGAGGTTGAGGGGAGAAATGCTTGAACCCAGAGAGTGTGCTGAGTGAATTGA 1715  
Db 129 GCGAGAGGTTGAGGGGAGAAATGCTTGAACCCAGAGAGTGTGCTGAGTGAATTGA 188

QY 1716 GCAACTGAATCCAGCCCTGGGCGCAGAGTGAAGTCTCTCAAAAAAAGGA 1775  
Db 189 GCAACTGAATCCAGCCCTGGGCGCAGAGTGAAGTCTCTCAAAAAAAGGA 247

QY 1776 TCGTCTCAACCTTTGCCCTCTACTGCAACATTTTGTATTTGAATGAAGTAACTTCC 1835  
Db 248 TCGTCTCAACCTTTGCCCTCTACTGCAACATTTTGTATTTGAATGAAGTAACTTCC 307

QY 1836 ATACTTATGCTGTTAATTAATCTTCAATCTCACTAGAGTGAAGCAATCTTCACTGCTT 1895  
Db 308 ATACTTATGCTGTTAATTAATCTTCAATCTCACTAGAGTGAAGCAATCTTCACTGCTT 367

QY 1896 CCTAATGGGAGATGCTTGGCCAGCAGGCTCTCACTGTGTGTACACAGAGAGCACTGA 1955  
Db 368 CCTAATGGGAGATGCTTGGCCAGCAGGCTCTCACTGTGTGTACACAGAGAGCACTGA 427

QY 1956 TCCAGTCAAGCCATACAGCTGTCCACACTGAAGAGTGTCTTACAAAGCCTGAATCA 2015  
Db 428 TCCAGTCAAGCCATACAGCTGTCCACACTGAAGAGTGTCTTACAAAGCCTGAATCA 487

QY 2016 AATGTTAGCTTAATGATTAATAATCCAGACTACT 2051  
Db 488 AATGTTAGCTTAATGATTAATAATCCAGACTACT 523

RESULT 14  
LOCUS BM707813 529 bp mRNA linear EST 28-FEB-2002

DEFINITION UI-E-C11-afa-1-23-0-UI.r1 UI-E-C11 Homo sapiens cDNA clone  
ACCESSION U1-E-C11-afa-1-23-0-UI 5', mRNA sequence.  
VERSION BM707813  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Bonaldo, M.F., Lennon, G. and Soares, M.B.  
AUTHORS Normalization and subtraction: two approaches to facilitate gene  
TITLE discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: Bento-soares@uiowa.edu  
tissue procurement: Dr. Gregg Hageman  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 435-519, >ALU  
Seq primer: M13 Reverse.  
Location/Qualifiers  
1. 529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-C11-afa-1-23-0-UI"  
/tissue\_type="RPE and Choroid"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/notes="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-C11 is a normalized cDNA library containing the  
following tissue(s): RPE and Choroid. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pRTT3-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is ACCTA.  
This library was created for the program, Gene Discovery  
in the Visual System, supported by National Eye Institute  
(NEI)."  
ORIGIN  
Query Match 22.2%; Score 492.6; DB 12; Length 529;  
Best Local Similarity 97.7%; Pred. No. 2.8e-106;  
Matches 509; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
QY 1101 CACTACAGAAAGGATGGCAGAACTGCTCACTCTAGAACTGTAGTGGTGAAGTT 1160  
DB 1 CACTACAGAAAGGATGGCAGAACTGCTCACTCTAGAACTGTAGTGGTGAAGTT 60  
QY 1161 GAGGGGTCTCAGAGCCCTTATGGTCACTTTTCTCACTGCTGGAACTCACCACCAATATCTT 1220  
DB 61 GAGGGGTCTCAGAGCCCTTATGGTCACTTTTCTCACTGCTGGAACTCACCACCAATATCTT 120  
QY 1221 CTGCTTCTCTGGGTCAGCCAAAGCTGTCACAAAATCATGATATTTCCCTTATTCAG 1280

DB 121 CTGCTTCTCTGGGTCAGCCAAAGCTGTCACAAAATCATGATATTTCCCTTATTCAG 180  
QY 1281 ATTTCCTGACACTGTCACTCACTCAATTTATTAACACCCCACTTCAAGCCCAATCAGCTGGAG 1340  
DB 181 ATTTCCTGACACTGTCACTCACTCAATTTATTAACACCCCACTTCAAGCCCAATCAGCTGGAG 240  
QY 1341 GAAGTGAATCTTCCCTTTTCTGGAATTTCTCAAGCAAGTACTTTCACGGGTGAGAACAGCA 1400  
DB 241 GAAGTGAATCTTCCCTTTTCTGGAATTTCTCAAGCAAGTACTTTCAC-GGTAGAACAGCA 299  
QY 1401 GCTATTATGATTTGAAAACCTTAAAGGCAACAATTTCACTTCTTCTTGAAGCTTAAGACA 1460  
DB 300 GCTATTATGATTTGAAAACCTTAAAGGCAACAATTTCACTTCTTCTTGAAGCTTAAGACA 359  
QY 1461 GGAACCTTGGCAAACTCTGTGGCTGTTCAAGCAAGATTTCAATTTATTAAGATCTTGT 1520  
DB 360 GGAACCTTGGCAAACTCTGTGGCTGTTCAAGCAAGATTTATTAAGATCTTGT 419  
QY 1521 CTGGGCTGGGTGGAGGCAAGTGAATCAAGAGGTGAGAGTTGAGACCACTGG 1580  
DB 420 CTGGGCTGGGTGGAGGCAAGTGAATCAAGAGGTGAGAGTTGAGACCACTGG 479  
QY 1581 CCAACATGATGAAACCCCATCTCTACCAAAAAAATATCAAA 1621  
DB 480 CCAACATGATGAAACCCCATCTCTACCAAAAAAATATCAAA 520  
RESULT 15  
AA307119  
LOCUS AA307119  
DEFINITION EST178031 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end,  
mRNA sequence.  
AA307119  
VERSION AA307119.1 GI:1959449  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 469)  
Adam, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,  
Bull, C.J., Lee, N.H., Kirkness, E.F., Weinrock, K.G., Gocayne, J.D.,  
White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Whi, C.,  
Clayton, R.A., Cline, T.R., Cotton, M.D., Batle-Hughes, J., Fine, L.D.,  
Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J., Geoghegan, N.S.,  
Glock, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,  
Kelley, J.M., Kelley, D.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M.,  
Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,  
Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R.,  
Small, K.V., Spriggs, T.A., Uterback, T.R., Weidman, J.F., Li, Y.,  
Bedarik, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, B.Y.,  
Dimke, D., Feng, D.-F., Ferris, A., Fischer, C., Hastings, G.A.,  
He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,  
Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H.,  
Raymond, L., Wei, Y.F., Wang, J., Xu, C., Yu, G.L., Ruben, S.M.,  
Dillon, P.O., Pannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,  
Fraser, C.M. and Venter, J.C.  
Initial assessment of human gene diversity and expression patterns  
based upon 83 million nucleotides of cDNA sequence  
Nature 377 (6547 Suppl), 3-174 (1995)  
JOURNAL 96026280  
MEDLINE 756098  
PUBMED  
COMMENT Other ESTs: THC123650  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene



Index (<http://www.tigr.org/tdb/hgi/hgi.html>)  
Seq primer: M13 Reverse.  
Location/Qualifiers

## FEATURES

source  
1..469  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="ATCC (inhost):108624"  
/db\_xref="taxon:9606"  
/tissue\_type="colon"  
/cell\_type="KM12C"  
/cell\_line="KM12C(HCC)-parental human colon  
carcinoma/Dukes B2"  
/clone\_lib="Colon carcinoma (HCC) cell line"  
/note="Organ: colon; Vector: pbluescript SK-; Site\_1:  
EcoRI; Site\_2: XhoI"

## ORIGIN

Query Match 21.0%; Score 466.4; DB 9; Length 469;  
Best Local Similarity 99.6%; Pred. No. 4.8e-100;  
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 50 CAATCAGAGAGAGAGAGATGCTCAGCGTGGCATCATTTGCGCGTTCTTAGGCTTGA 109  
DB 1 CAATCAGAGAGAGAGAGATGCTCAGCGTGGCATCATTTGCGCGTTCTTAGGCTTGA 60  
QY 110 GTCCCATGATCAACATCTCCAGGGCAAACTCAAGGACCAACTACTGTGGCCCAAG 169  
DB 61 GTCCCATGATCAACATCTCCAGGGCAAACTCAAGGACCAACTACTGTGGCCCAAG 120  
QY 170 GGAATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGACGCAAAAGCTTAG 229  
DB 121 GGAATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGACGCAAAAGCTTAG 180  
QY 230 GGGCCAGAGAGACCAAGAGGCTTGAGCTTAAAGCTTGAGCGCTTCAAGTGTGCCCC 289  
DB 181 GGGCCAGAGAGACCAAGAGGCTTGAGCTTAAAGCTTGAGCGCTTCAAGTGTGCCCC 240  
QY 290 ACTGTATCAGAGGCCAGGCTACTACAGTGCCCCACAGACGCCCTCAGCCCACTCCCAT 349  
DB 241 ACTGTATCAGAGGCCAGGCTACTACAGTGCCCCACAGACGCCCTCAGCCCACTCCCAT 300  
QY 350 GTTCTTCCCGCTAGAACCATCAGAGCGCGTCAAAAGCTTCAAGTGTCAAGGCATAGAC 409  
DB 301 GTTCTTCCCGCTAGAACCATCAGAGCGCGTCAAAAGCTTCAAGTGTCAAGGCATAGAC 360  
QY 410 CAAAGACAAAAGCTTAAAGACTGTAGTCTGAGGCGCAAGAAAAGTTTGAATTGCTCTC 469  
DB 361 CAAAGACAAAAGCTTAAAGACTGTAGTCTGAGGCGCAAGAAAAGTTTGAATTGCTCTC 420  
QY 470 AGAGAGGATGGGGCTTGAATGAGCAACCAAGATATCTCAAGTGAGG 518  
DB 421 AGAGAGGATGGGGCTTGAATGAGCAACCAAGATATCTCAAGTGAGG 469

Search completed: March 25, 2004, 07:21:33  
Job time : 5679 secs